# Order Exposure in High Frequency Markets*

Bidisha Chakrabarty, Terrence Hendershott, Samarpan Nawn, and Roberto Pascual†

May, 2019

**Abstract**

We show that high frequency traders (HFTs) make extensive use of hidden limit orders. These orders are of small share sizes, aggressively placed near the best quotes, and generally not informationally motivated. Theory suggests that traders place hidden orders to limit their option value or delay information exposure. Our tests show that neither the free-option theory nor the information-revelation theory can explain HFTs' hidden order usage patterns. To provide groundwork for new theory, we offer and empirically validate two conjectures about HFTs' motives for order non-exposure.

Keywords: Hidden orders, high frequency trading, order exposure

*JEL Classification*: G11; G12; G14; G15, G24

†: bidisha.chakrabarty@slu.edu, hender@haas.berkeley.edu, samarpan.nawn@iimu.ac.in, rpascual@uib.es

We thank David Abad, Michael Brolley, Darrell Duffie, Bart Frijns, Corey Garriott, Sudip Gupta, Frank Hatheway, Vitto Mollica, Marta Khomyn, Venky Panchapagesan, Satchit Sagade, Ajay Shah, Andriy Shkilko, Irena Vodenska, Pradeep Yadav and seminar participants at the Auckland Finance Meeting (Queenstown), Auckland University of Technology, Australasian Finance and Banking Conference (Sydney), Emerging Markets Finance Conference (IGIDR Mumbai), Financial Management Association European Conference (Kristiansand), Financial Market Liquidity Conference (Budapest), Northern Finance Association Meeting (Charlevoix), NSE-NYU Conference, Saint Louis University, Swiss Society for Financial Markets Research (Zurich), Universidad CEU Cardenal Herrera, Wilfrid Laurier University, University of Paris-Dauphine, Stockholm Business School, and European Financial Management Association Meeting (Ponta Delgada), for useful comments.

* This working paper is part of the NSE-NYU Stern School of Business Initiative for the Study of the Indian Capital Markets. The authors acknowledge the support of the initiative. The views expressed in this Working Paper are those of the authors and do not necessarily represent those of the National Stock Exchange of India (NSE) or New York University (NYU). Pascual acknowledges the financial support of the Spanish DGICYT project ECO2017-86903-P, and the Generalitat Valenciana Grant Prometeo/2017/158. Chakrabarty acknowledges the Edward Jones endowment. This paper was partially done while Pascual was visiting the Auckland Center for Financial Research (ACFR) of the Auckland University of Technology.



Electronic copy available at: https://ssrn.com/abstract=3074049

## 1. Introduction

Public stock exchanges are not fully transparent; opacity – the choice to hide orders – is on the rise in financial markets. The SEC's market data center shows that hidden volume's contribution to trades increased from 15% to over 30% in the US between 2012 and 2017[1]; our own estimates (see Figure 1) show that 30% (50%) of the Nasdaq limit order book near the best quotes is hidden on low (high) volatility days.[2]

**[Figure 1]**

A parallel trend during the last decades has been the rise of high frequency trading (HFT). According to the Tabb Group, while HFT accounted for 20% of U.S. equity volume in 2005, in 2016 it had reached 50%. While correlation is not causation, the outsized influence of HFT in modern financial markets raises the question: Do high-frequency traders (HFTs) contribute to opacity in public exchanges by hiding their trading interest? We develop testable propositions from theories of order exposure to address this question and use multi-country data that flag trader and order types to test these propositions. We find that HFTs extensively use hidden orders as part of their trading strategy. This finding is surprising because the logic of the extant theories of hidden order usage suggests that HFTs should have little incentive to hide limit orders.

Why should a trader choose to hide her trading interest? Extant theory models this choice depending on whether the trader is informed (Moinas, 2010; Boulatov and George, 2013) or uninformed (Buti and Rindi, 2013) which ignores differences in latency (speed), quotation frequency, or monitoring intensity among traders. Most empirical tests that address these theories are done in markets of the pre-HFT era (De Winne and D'Hondt, 2007; Bessembinder, Panayides and Venkataraman, 2009 (hereafter, BPV); Pardo and Pascual, 2012), where speed was not as important an issue as in modern markets. Therefore, there is a need to extrapolate the theoretical rationale for order non-exposure to markets populated by HFTs.

One strand of the literature we call the *free-option theory* (Buti and Rindi, 2013) focuses on an uninformed liquidity provider who, by displaying large order sizes, exposes herself to the risk of being picked off by faster traders, adversely selected by

---

[1] https://www.sec.gov/marketstructure/datavis/ma_exchange_hiddenvolume.html#.XM-0auhKg2w

[2] Opacity in financial markets come from orders that are fully or partly hidden (iceberg) in lit exchanges as well as venues known as dark pools that are completely opaque. In lit markets (the focus of this study), fully hidden orders are more prevalent in North American markets (e.g., US, Canada) while iceberg orders are widespread in Europe and the Asia-Pacific (e.g., Spain, France, India, Australia).



Electronic copy available at: https://ssrn.com/abstract=3074049

informed traders, or undercut by parasitic traders. In this framework, the uninformed trader hides her orders to mitigate their option value (Copeland and Galai, 1983). Should this narrative apply to HFTs? The literature shows that HFTs are a significant source of liquidity supply (Hagströmer and Nordén, 2013), but they use smaller order sizes (O'Hara, 2015), and monitor markets in near-continuous time (Hoffmann, 2014), resulting in high rates of ultra-fast cancellations (Hasbrouck and Saar, 2009). Their limit orders should therefore have a low option value. Moreover, hidden orders lose time priority per exchange rules, which increases their time to execution. Since the success of HFTs' trading strategies relies on speed (Baron, Brogaard, Hagströmer, and Kirilenko, 2018), they should be better off displaying their orders and quickly canceling or updating their quotes as market conditions necessitate.

A second branch of the literature models informed traders' motives for hiding orders. The *information-revelation theory* posits that informed traders may use hidden orders to obscure their trading intentions (Moinas, 2010), thereby reducing the expropriation of informational rents (Boulatov and George, 2013). Studies show that HFTs' trades (Brogaard, Hendershott, and Riordan, 2014) and orders (Chordia, Green, and Kottimukkalur, 2018) carry information, although Weller (2018) emphasizes that HFTs are not informed in the traditional sense of producing new information. Rather, they contribute to price discovery by rapidly incorporating signals gleaned from order flow (Hirschey, 2018; Korajczyk and Murphy, 2019) or public news (Chakrabarty, Moulton, and Wang, 2019). In this case we expect HFTs to use displayed orders since such information is short lived and, by losing time priority, hidden orders delay execution.

Thus, whether HFTs are informed or uninformed, given the trading technology they deploy and the unique features of hidden orders, extant theory suggests that HFTs should display their orders.

To test if that, indeed, is the case, we need data that (a) flag HFT versus other traders, and (b) provide order level information including the display condition (hidden or not). Publicly available trade and quote data generally do not have either flag. We use proprietary data from two markets that provide such identifiers. Our primary data come from the National Stock Exchange of India (NSE), the fifth largest market in the world



Electronic copy available at: https://ssrn.com/abstract=3074049

by number of trades.[3] The NSE data furnish rich details on trader accounts, using which we classify each order as coming from one of three mutually exclusive trader types: proprietary algorithmic traders (i.e., HFTs), other ("agency") algorithmic traders (AATs), and non-algorithmic traders (NATs). The NSE allows iceberg orders and in these data we can identify both the displayed and the hidden portions of each order. Our second data source is the Nasdaq exchange in the US which allows fully hidden orders. The Nasdaq data provide one-minute snapshots of the ten best bids and offers in the order book. For each snapshot, we see all standing limit orders, whether they are hidden or displayed, and whether they were placed by HFTs or non-HFTs. We use the term hidden limit orders (HLOs) for both fully hidden and iceberg orders, noting that the NSE HLOs are iceberg orders while the Nasdaq HLOs are fully hidden.

We answer three primary questions related to HFTs' hidden order usage – the "whether", the "how", and the "why." In response to whether HFTs use hidden orders, we find that they make extensive use of HLOs. In the NSE, in large-cap firms 10.38% (9.83%) of all limit orders (share volume) submitted by HFTs are HLOs. Corresponding numbers for mid-cap and small-cap firms are 36.0% (34.42%) and 15.84% (15.23%), respectively. In the Nasdaq, HFTs hide 21.8% (15.25%) of all limit orders (share volume) in large-cap stocks, 23.17% (34.71%) for mid-cap stocks, and 31.65% (47.84%) for small-cap stocks.

Analyzing order placement in different layers of the book, for the NSE we find that HFTs place 46.03% (1.5%) of their hidden (displayed) orders in large stocks at or better than the best quotes. In fact, over 97.72% of HFT's HLOs in large stocks are placed within the first five ticks from the best quotes. In contrast, NATs place 39.12% of their HLOs away from the five best ticks. In small stocks, HFTs' HLOs are rarely placed away from the five best ticks while NATs place the bulk of their HLOs far away from the best quotes. The Nasdaq data corroborate that HFTs' HLOs are more aggressive than their displayed limit orders (DLOs) as well as the HLOs of non-HFTs.

So data from both markets indicate that HFTs use HLOs. But how efficiently do they use these orders? We model this part of the investigation on BPV who find that HLOs have a lower probability of completion and take longer to execute compared to similar DLOs, although DLOs have a higher implementation cost. How do HFTs

[3] In our sample period, HFTs contribute 33% of the total daily volume on the NSE. See https://www.nseindia.com/research/content/1314_BS6.pdf



Electronic copy available at: https://ssrn.com/abstract=3074049

manage this cost-benefit trade-off vis-à-vis other traders? To test the effectiveness of HFTs' hidden order usage, we model the execution probabilities of HLOs placed by HFTs versus other traders. We find that HFTs' HLOs have the *highest* likelihood of execution. Although HLOs lose time priority, HFTs' hidden orders have a similar (higher) fill rate than their displayed orders for large (mid or small) caps, suggesting that HFTs strategically place HLOs in anticipation of short-term volatility increases, which increases the likelihood of execution. We also model the time to full execution of HLOs vis-à-vis other orders using survival analysis, as in Lo, MacKinlay, and Zhang (2002). This test is particularly relevant in our context, since iceberg orders may mechanically induce a protracted time to completion. Results show that although compared to DLOs, HLOs take longer to fully execute, HLOs placed by HFTs execute faster than those placed by other traders.

Clearly HFTs benefit from the increased likelihood of execution and reduced time to completion of their HLOs. But any benefit must be weighed against the cost incurred. To estimate the costs, we use the implementation shortfall metric (Perold, 1988). This metric has two components: effective cost (price impact), and the opportunity cost of non-execution (which measures forgone profits). We find that HFTs face higher effective cost for hidden orders, which is expected since HFTs use more aggressive HLOs. However, their opportunity cost of non-execution is lower, indicating less adverse price movements after their hidden order submissions. When combined, the lower opportunity costs either compensate for, or exceed, the higher execution costs and overall HFTs' HLOs have a lower implementation shortfall. These findings suggest that HFTs use HLOs more efficiently than non-HFTs.

Our final set of tests address the why question. First we test the free-option theory which suggests that large limit orders are more likely to be hidden. Do HFTs hide (relatively) larger limit orders? Our results suggest that is not the case. In fact, HFTs use smaller share sizes for HLOs. In the NSE, HFTs' HLOs average 456.58 shares compared to 1139.59 shares for NATs. For displayed orders, the patterns reverse: HFTs use larger DLOs (1150.50 shares) than NATs (309.27 shares). 76.28% (5.11%) of HFTs' HLOs (DLOs) in large firms are placed in the under-50-shares category while for mid and small firms, this rises to 98.72% and 83.96%, respectively. We also estimate the probability of hiding a limit order conditional on order size and find that HFTs are more likely to hide smaller orders. These patterns are also present in the Nasdaq data.



Electronic copy available at: https://ssrn.com/abstract=3074049

Thus, our results find no confirmation for the free-option theory when extended to the use of HLOs by HFTs.

To test the information-revelation theory, we examine whether HFTs' HLOs are informationally motivated using three complementary metrics. First, we measure the average information content of HLOs for each trader type. Second, we decompose the order-flow related component of the efficient price variance into proportions attributable to each trader-type (HFT, AAT, NAT) – order-type (hidden, displayed) combination. Third, we measure the information share (Hasbrouck, 1995) of each trader-type order-type combination. We find that HFTs' HLOs have an insignificant price impact once we account for order aggressiveness, they explain the smallest portion of order-flow related efficient price volatility, and they have the lowest information share of all trader-type order-type combinations. Overall, our findings are inconsistent with HFTs using HLOs to trade on time sensitive information and fail to confirm the information-revelation theory.

Collectively, these results indicate that HFTs use HLOs neither to manage free-option risk nor to manage information revelation. Existing models, therefore, do not explain why HFTs should use hidden orders, which calls for new theory to explicitly model HFTs' order exposure choice. To that end, and as a first step, we empirically investigate two possible reasons why HFTs may use HLOs: (a) to undercut standing quotes and compete to supply liquidity,[4] and (b) in anticipation of peaks in short term volatility. We note here that this is not an exhaustive list of the possible reasons why HFTs use HLOs, but rather tests based on some characteristics of HFT strategies documented in contemporary studies.

HFTs' ultra-fast algorithms put them in a position to anticipate other traders' orders (Hirschey, 2018) or detect institutional investors' orders that use order-splitting algorithms (van Kervel and Menkveld, 2019). Using HLOs, HFTs could undercut standing orders without revealing their presence. Additionally, there could be some speed advantage to letting the exchange's engine (software) reveal each successive tranche of an iceberg order, rather than transmitting several small DLOs from the HFTs'

---

[4] Offering minimal price improvement to undercut standing quotes and move up in the order queue may enhance liquidity supply and narrow the bid-ask spread, or adversely impact other liquidity suppliers if such quotes are used to persistently jump ahead of standing orders. Since HFTs do not have any fiduciary obligation towards the traders whose quotes they undercut, our tests do not address the illegal practice of "front running," where the undercutting party has such obligation to the party whose orders are undercut.



Electronic copy available at: https://ssrn.com/abstract=3074049

server by monitoring market conditions. We define an undercutting order as a limit order that (i) is placed immediately after another submission on the same side of the market, (ii) comes in under 10 milliseconds of the previous order, and (iii) improves upon the previous price. We find that HFTs are more likely to use HLOs than DLOs to undercut existing orders at or near the best quotes using aggressively priced HLOs.

Foucault, Hombert and Roşu (2016) show that traders with speed advantages deploy anticipatory trading strategies, and evidence in Hirschey (2018) confirm that HFTs anticipate order flow. Extending this line of reasoning, we ask if HFTs also anticipate volatility and place non-aggressive HLOs in periods before volatility peaks, thereby improving their probability of execution (a result we documented earlier). Our empirical tests also confirm this conjecture about the use of HLOs by HFTs.

This study sits at the cusp of two important issues facing investors and regulators – market opacity and high-speed trading. Research shows that when markets allow traders the facility to hide orders, they substitute non-displayed for displayed orders and change their trading aggressiveness (Bloomfield, O'Hara, and Saar, 2015). Meanwhile, improved (pre-trade) transparency can increase liquidity and the informational efficiency of prices (Boehmer, Saar, and Yu, 2005). Since transparency is a cornerstone of the SEC's investor protection function, current trends have regulators worried that opacity may be attractive to "bad-actors" (see SEC Chairman's speech).[5] The growth of HFT has also been accompanied by a frenzy of media commentaries on its inherent unfairness. Although studies find that HFT has both positive (Brogaard, Hendershott, and Riordan, 2014) and negative (Budish, Cramton, and Shim, 2015) effects, there has been no evidence to date linking HFTs to market opacity.

To our knowledge this is the first study to document that HFTs make extensive use of HLOs in lit markets. These orders are different in characteristics (e.g., size, aggressiveness), information content (e.g., contribution to price variance), and usage (e.g., liquidity supply, undercutting) than the HLOs of non-HFTs, and do not fit the logic of order exposure modeled in extant theory. These results are robust in that they hold for both consolidated (NSE) and fragmented (US) markets, and for iceberg as well as fully hidden orders, allowing us to rule out market design or the choice between partial versus full non-exposure as explanatory factors.

[5] https://www.sec.gov/news/speech/speech-clayton-2017-11-08



Electronic copy available at: https://ssrn.com/abstract=3074049

We structure the paper as follows. Section 2 reviews the free-option and the information-revelation theories. Section 3 presents the institutional details of the NSE market, identification of trader account types and HLOs, and the samples for the NSE and the Nasdaq markets. Section 4 addresses the "whether" question using descriptive statistics and tests of HLO use by trader types. Section 5 examines the efficiency of HLO use by HFTs by modeling the likelihood, time to execution, and implementation costs, thereby answering the "how" question. Section 6 evaluates whether the free-option theory and/or the information-revelation theory explain HFTs' use of HLOs and, finding confirmation for neither, in Section 7 we test two conjectures for HFTs' HLO use – namely undercutting and volatility anticipation. These two sections together address the "why" question. Section 8 concludes. Results from the Nasdaq sample, book-building from the NSE data, variable description, additional tests, and selected test procedures are presented in the accompanying Internet Appendix.

## 2. Theoretical rationale for the use of hidden orders

Extant theory supports two competing rationales for the use of HLOs in lit markets: the *free-option theory* and the *information-revelation theory*.

### 2.1 The free-option theory

Limit orders to buy (sell) are free put (call) options in that a trader who submits a limit order on the bid (ask) side of the book writes an option for the counterparty to sell (buy) a limited amount of shares at a pre-determined limit price (strike), but receives nothing (no premium) in exchange (Copeland and Galai, 1983). As a result, trading through limit orders is costly since these free options can be executed in the money. This happens either when the limit order becomes mispriced after the unexpected arrival of adverse public news and is picked off by a faster trader, or when it is adversely selected by a better-informed trader. Building on option pricing theory, the free-option risk of a limit order (its likelihood of executing in the money) increases with factors such as the closeness of its limit price to the quote midpoint (aggressiveness), its size, the expected time it will stand in the limit order book (LOB), and the volatility of the stock price.

To mitigate this risk, limit order traders can monitor their orders and cancel or revise them as needed, but monitoring is costly (Liu, 2009; Fong and Liu, 2010). Or, they can follow a passive strategy and place their limit orders away from the best quotes, but that



Electronic copy available at: https://ssrn.com/abstract=3074049

increases their non-execution risk which involves an opportunity cost in the form of forgone profits (Hasbrouck and Saar, 2009). Alternatively, investors can mitigate the free-option risk by (totally or partially) hiding their limit orders (Aitken, Berkman, and Mak, 2001). By choosing to hide, the investor reduces the option value of the limit order to both better informed and faster traders, while still placing the order close to the best quotes.

Traders who display large trading interests using limit orders also expose themselves to parasitic traders such as front-runners (also known as quote-matchers or penny jumpers), who profit by trading ahead of large limit orders they expect will have significant price impacts (Harris, 1997). Buti and Rindi (2013) develop a model where large uninformed limit order traders face exposure costs in the form of aggressive undercutting by parasitic traders. In their model, HLOs prevent the friction generated by such undercutting.

### 2.2. The information-revelation theory

Early market microstructure theory (e.g., Glosten and Milgrom, 1985) presumed that informed traders buy and sell aggressively using market orders. More recently, theoretical (e.g., Kaniel and Liu, 2006), empirical (e.g., Anand, Chakravarty, and Martell, 2005), and experimental (e.g., Bloomfield, O'Hara, and Saar, 2005) studies suggest that informed traders may choose to provide rather than take liquidity, thereby using limit orders. O'Hara (2015) claims that sophisticated informed traders rarely cross the spread in modern high frequency markets. Consistently, Brogaard, Hendershott, and Riordan (2019) find that nowadays price discovery occurs predominantly through limit orders.

Our question is whether informed traders hide their limit orders. The information-revelation theory suggests that informed traders may use HLOs to obscure their trading intentions, so as to delay the revelation of their private information, minimize the price impact of their trades, and thus maximize the rents they can extract from their private signals (Harris, 1997). The notion that informed traders look for ways to camouflage trading is not new. Admati and Pfleiderer (1988), for example, propose a model in which informed traders conceal their market order trading among that of the uninformed. Similarly, HLOs provide an opportunity for informed traders to camouflage their trading.



Electronic copy available at: https://ssrn.com/abstract=3074049

Intuitively, if the private information is substantial and will not become public soon, informed traders may choose to trade less aggressively (e.g., Harris, 1998; Kaniel and Liu, 2006). In such a case, they may choose HLOs to transact larger volumes without signaling their presence. Moinas (2010) develops a model where liquidity providers, who are informed with some probability, may choose to hide their large limit orders so as to obscure their presence to uninformed liquidity takers who may otherwise refuse to trade, especially when information asymmetry risk is high.

Boulatov and George (2013) examine whether HLOs enhance market quality by attracting traders that otherwise would not provide liquidity in lit markets or, rather, degrade market quality by leveraging the advantage of informed traders over the uninformed traders. In their model informed traders jointly decide whether to make or take liquidity, and whether their liquidity supply is hidden or displayed. The model predicts that allowing HLOs will draw more informed traders into liquidity provision, intensifying the competition for liquidity supply. As a result, market quality improves.

In sum, the free option theory posits that HLOs are used by uninformed traders to manage the option value of their (large) limit orders, while the information revelation theory posits that traders who possess information that can be appropriated by faster traders will use HLOs to minimize such exposure risks.

## 3. Institutional features, trader identification, and sample selection

### 3.1. NSE: trading protocol, iceberg orders, and HFT identification

With over 80% of the total traded volume, the NSE is the dominant market for its 1300+ listed stocks. It is a completely order driven market. The NSE allows traders to place dark orders by choosing the “iceberg” option with a mandatory minimum exposure of 10% (of the original volume). Once the first tranche is executed, the next tranche is automatically displayed. All tranches are of the same size (10% or greater of the original order). The market operates on price-exposure-time priority whereby non-displayed volume loses time priority to any displayed volume at the same price. Thus, the iceberg order provision of the NSE is similar to that used on the Euronext and analyzed in BPV. There are no dark pools in this market, so traders who want to hide orders have only iceberg orders in the lit market as their option. For more details on the general institutional features of this market, we refer readers to Kahraman and Tookes (2017).



Electronic copy available at: https://ssrn.com/abstract=3074049

We obtain order and trade data directly from the NSE. For each trading day, we access a message file and a trade file. The message file contains every message for each stock including the ticker symbol, price, quantity and timestamp in jiffies (one jiffy is 33.3564 picoseconds or $(1/2\text{^}16)^{th}$ of a second). Similar to the ITCH data of the Nasdaq platform, for every order this file includes order entry, modification, execution and cancellation events. The trade file contains analogous information for each trade. By allowing temporal tracking of each order and matching orders to trades, these data allow us to build the complete limit order book (LOB) at any time instant. Internet Appendix A provides details on building the LOB from the NSE data.

Both the message and the trade files provide several flags or identifiers. For the purposes of this study, we use three of these flags: *Client*, *Order Entry Mode*, and *Modifier condition*. *Client* classifies trader accounts into *Custodian*, *Proprietary* and *Others*. *Custodian* represents traders who are members of the exchange but do not conduct their own clearing or settlement. This group comprises primarily of foreign institutional investors, mutual funds, and financial institutions. The *Proprietary* flag applies to members of the exchange who trade for their own proprietary accounts, and *Other* applies to all other customers of the exchange who employ their own clearing member.

*Order Entry Mode* flag applies to each *Client* flag and shows one of the two possible order entry systems used to interact with the NSE's limit order market: *Algorithmic* if order entry and management is done using an algorithm and *Non-Algorithmic* if a client uses a manual system. The product of the three *Client* flags and two *Order Entry Modes* enables us to identify six distinct message originations. Our particular focus is on the *Proprietary* client using *Algorithmic* order entry mode to trade on their own account. That is the definition of HFTs (SEC, 2010[6]) and we can cleanly identify the message traffic from HFTs in our data. We group all other messages with the *Algorithmic* flag into the agency algorithmic trader (AAT) type and all messages flagged with *Non-Algorithmic* order entry mode as non-algorithmic traders (NATs).

A key advantage of our identification is that, unlike previous studies, we classify HFT at a finer granularity. For example, when a trader conducts prop trades using algorithms, we classify those trades as HFT, but when this same trader conducts client

[6] Securities and Exchange Commission concept release on equity market structure. Available at https://www.sec.gov/rules/concept/2010/34-61358.pdf



Electronic copy available at: https://ssrn.com/abstract=3074049

trades using algorithms, we do not count those as HFT. This overcomes some known limitations of popular HFT identifying databases that club all HFTs as pure-play (e.g., the Nasdaq HFT database used in Brogaard et al., 2014) or allow for mixed categories that cannot be exactly classified as HFT (e.g., the EUROFIDAI data used in Boussetta, Lescourret, and Moinas, 2017).

Finally, the *Modifier* flag identifies all orders entered with an iceberg condition and shows the minimum display volume, allowing us to see both the lit and dark proportions of each iceberg order.

### 3.2. Nasdaq hidden orders and trader identification

The Nasdaq data of this study has been used in several recent papers, for example Carrion (2013), and O'Hara, Yao, and Ye (2014). But unlike these studies that use the trade files, we use the LOB files provided in these data. Nasdaq allows traders to fully hide an order. In our data, all available liquidity on the Nasdaq book is shown at one-minute snapshots from 9:30 a.m. to 4:00 p.m. (inclusive). We have data for the first full week of the first month of each quarter during our sample period (2008 and 2009) the crisis week of Sept 15 – 19, 2008, and the week of Feb 22 – 26, 2010. These records have a buy/sell indicator to denote whether liquidity was on the bid or offer side, limit price, a flag to indicate HFT, a flag to denote if the liquidity is displayed or hidden, the ticker symbol, order size, book snapshot time and date. For each snapshot, order by order records representing the ten best price levels (displayed and hidden) on each side of the market are shown.

### 3.3 Sample selection

HFTs have a preference to trade large stocks (Brogaard et al. 2014). To ensure even representation of both HFTs and other trader types, from the NSE data we select a (market cap) stratified sample of 100 stocks as follows. We begin with the 1254 listed stocks in the NSE in September 2013, filter out 286 stocks that are not in continuous trading session in our sample period October to December 2013 (61 trading days). We also exclude firms that (i) have a closing price of Rs. 1 or lower, (ii) have fewer than 100 trades per day on average, (iii) trade less than 1000 shares a day, (iv) have a traded value per day of less than Rs. 100000 over the sample period, (v) have market-cap values in the Bloomberg and CMIE Prowess databases that diverge by over 10%, (vi) are involved in NSE or MSCI index changes. These filters reduce our universe of stocks



Electronic copy available at: https://ssrn.com/abstract=3074049

to 695, which we sort by market capitalization and group into deciles. From each decile, we select the top 10 stocks to generate the sample of 100, with 30 large-cap stocks, 40 mid-cap stocks and 30 small-cap stocks. All company information come from the CMIE Prowess (analogous to Compustat), a database of Indian firms which covers approximately 80% of the NSE stocks (Kahraman and Tookes, 2017). Panel A of Table 1 shows the descriptive statistics of our sample.

**[Table 1]**

The average firm in our sample has over 448 billion rupees market capitalization (about 7 billion USD per the exchange rate on 06/2017). Large-cap firms have a market capitalization of about 1465 billion rupees (22 billion USD), which is smaller than the large cap firms in the Nasdaq HFT dataset (Brogaard et al., 2014). Volume and number of trades are higher, and relative spread (ratio of the quoted spread to the quote midpoint) is much smaller for the large firms than mid-sized and the small firms, as expected. While both the accumulated displayed and hidden depths in the LOB are higher for large firms than mid- and small-sized firms, the differences are larger for displayed than for hidden depth.

To benchmark our direct identification of HFTs against much of the literature that uses proxies for HFT activity, in Panel B of Table 1 we report message traffic and cancellation statistics by trader categories and across the three market cap groups. Comparing across each row, we see that HFTs account for much greater message traffic (defined as the sum of submissions, cancellations, and revisions) either than the AATs or the NATs in the large cap stocks, but not in the mid-sized or the small stocks. However, when we scale message traffic by the number of trades executed, HFTs show a bigger presence even in the mid- and small-cap firms. This preponderance of HFTs to generate large message traffic volume echoes similar findings from the US equity markets (e.g., Hendershott, Jones, and Menkveld, 2011).

The Nasdaq dataset consists of trades and quotes for a sample of 120 stocks, stratified by market capitalization and evenly split between Nasdaq and NYSE listing. Table IB 1 in the Internet Appendix B shows some descriptive statistics of this sample, including HFTs' presence at the top of the order book and their hidden order usage. For the full sample, over 71% of the time there is hidden volume at the best quotes; this number goes up to almost 80% for large stocks and is about 67% for small cap stocks.



Electronic copy available at: https://ssrn.com/abstract=3074049

HFTs are at the best quotes 67.41% of time in the full sample and 93.23% (46.11%) in the large (small) cap stocks.

## 4. The "whether" question: Hidden order use

Both the theories we reviewed earlier suggest that we should not expect HFTs to hide their limit orders. In this section, we examine whether HFTs indeed use any HLOs. To do so, we begin by examining the placement of hidden and displayed orders in the LOB, and compute the accumulated displayed and non-displayed depth, both in the number of orders and in share volume. Table 2 reports the results.

**[Table 2]**

In Panel A, we show the proportion of HLOs relative to all limit orders submitted, both for the number of orders and the volume of shares. Comparing across the first row, 10.38% (9.83%) of all orders (volume) submitted by HFTs in large cap stocks are HLOs. Although HFT message traffic is highest in the large cap stocks (Panel B of Table 1), we find that HFTs' use of HLOs is greater for mid-cap stocks. They place 36% (34.42%) of all orders (share volume) as HLOs in mid-cap stocks. In Panel B, we show each trader type's share of both DLOs and HLOs. HFTs account for 34.67% of DLOs but only 9.28% of HLOs in the large stocks. In Table IB 2, we report statistics on hidden volume used by HFTs in the Nasdaq market. Paralleling the findings for the NSE, HFTs hide 21.8% (15.25%) of orders (volume) for large caps, 23.17% (34.71%) for mid cap, and 31.65% (47.84%) for small cap stocks. HFTs account for 44% of all HLOs placed in the LOB for our Nasdaq sample.

Position in the limit order queue is valuable (Lo et al., 2002; Moallemi and Yuan, 2017), especially for HFTs, whose profits depend on being the fastest. Therefore, we examine where in the LOB HFTs place their hidden and displayed orders. We build the order book at every order submission time and identify the position of order placement at four positions – price improving or better than the standing best bid and ask quotes ("Better"), the best bid and ask ("At"), up to the first five ticks from the best bid and ask ("Near") and the rest of the book ("Far"). Table 3 presents hidden and displayed order placement by the different trader categories for the three firm size groups.

**[Table 3]**



Electronic copy available at: https://ssrn.com/abstract=3074049

Comparing corresponding cells in Panels A and B, we find that while 25.11% of HFTs hidden orders in large stocks are placed at "Better" than the best quotes, less than 0.5% (0.47% in Panel B) of their displayed orders are price-improving. Within Panel A, we find that while 97.72% (25.11% + 20.92% + 51.68%) of HFTs' HLOs in large stocks are within the five best ticks, the comparable fraction for NATs is 65.70%. In fact, in all three firm size groups, HFTs place a greater proportion of HLOs at or better than the best quotes. For small stocks, HFTs rarely place any HLOs away from the five best ticks. NATs show the exact opposite pattern, placing the bulk of their HLOs far away from the best quotes.

For displayed order placement (Panel B) we find the opposite pattern. Both HFTs and NATs place a bigger proportion of their DLOs away from the best quotes. While HFTs use both the near and far regions of the LOB to place DLOs, NATs concentrate their displayed orders mostly far from the best quotes. These results are mirrored by the share volume placement. Non-parametric tests show that the difference in HLO and DLO use is significant for all three trader categories.

The Nasdaq data mirror similar patterns. Table IB III shows that HFTs place HLOs more aggressively at or near the best quotes compared to non-HFTs; this reverses when contrasting their DLOs vis-à-vis non-HFTs' DLOs. Overall, the results from both the NSE and Nasdaq data establish that hidden order usage by HFTs is extensive, greater or similar to that of other traders, which is in contrast to what theory suggests. They place these orders more aggressively than other traders, and closer to the best quotes.

## 5. The "how" question: Efficiency of order exposure

BPV show that order exposure entails a trade-off: higher likelihood and shorter time to execution for DLOs versus lower implementation cost for HLOs. However, their sample period is April 2003, when HFT activity was minimal, if at all present. In this section, we test how efficiently HFTs manage the trade-off in their hidden order usage.

First, we examine the likelihood of HLO execution, specifically contrasting HFTs with other traders, by estimating an ordered Logit model, as in Ranaldo (2004) and Pascual and Veredas (2010).[7] The regression equation is:

[7] For this and all following analyses, the estimation sample consists of data for December 2013, and only includes the 30 largest stocks in our full sample (in which HFTs are reasonably active) to ensure adequate



Electronic copy available at: https://ssrn.com/abstract=3074049

$$
\begin{aligned}
EXEC_{ij} = {} & \alpha_0 + \alpha_1 Aggr_{ij} + \alpha_2 OrdSize_{ij} + \alpha_3 HLO_{ij} + \alpha_4 HFT_{ij} + \alpha_5 AAT_{ij} + \\
& + \alpha_6 HLOHFT_{ij} + \alpha_7 HLOAAT_{ij} + \alpha_8 RSpr_{ij} + \alpha_9 DepthSame_{ij} + \\
& + \alpha_{10} DepthOpp_{ij} + \alpha_{11} LOBImb_{ij} + \alpha_{12} LastHalfHour_{ij} + \alpha_{13} OI_{ij} + \\
& + \alpha_{14} TrdFreq_{ij} + \alpha_{15} Mom_{ij} + \alpha_{16} Volat_{ij} + e_{ij}
\end{aligned} \qquad [1]
$$

where the dependent variable ($EXEC_{ij}$) is an ordinal variable that signifies the degree of completion of the limit order $j$ for stock $i$. It takes three possible values: $EXEC_{ij}$ = 1 indicates that the limit order is cancelled before execution, $EXEC_{ij}$ = 2 indicates that the limit order is partially executed and then cancelled, $EXEC_{ij}$ = 3 indicates that the limit order is fully executed. We exclude market and marketable limit orders and drop fleeting orders (Hasbrouck and Saar, 2009), because they are not intended to be executed. Revisions of non-executed orders are treated as the same order while revisions of partially-executed orders are treated as new submissions. The variables of interest are the dummies $HLO_{ij}$ (which indicates if an order is hidden or not), $HFT_{ij}$ (indicates if an order is submitted by an HFT or not), and the interaction term $HLOHFT_{ij}$. We use NATs as the residual trader type.

Control variables are motivated from the previous literature (De Winne and D'Hondt, 2007; BPV). These variables include order characteristics (such as aggressiveness, and total order size), limit order book characteristics (such as the relative bid-ask spread, depth on the same and opposite sides, and LOB imbalance), and market conditions (such as a dummy for the last half hour of the trading session, order imbalance, trade frequency, stock momentum, and volatility).[8] For comparability across stocks, we normalize order size and trade size by dividing the actual observations by the stock's average daily trading volume. Internet Appendix C lists the definitions of variables.

The model is estimated on a stock-by-stock basis with the *t*-statistic for testing the significance of each variable computed using the Chordia, Roll, and Subrahmanyam (2005) method.[9] Table 4 reports the results.

---

number of observations for the models to converge. In this subsample, HLOs represent 15% of the total volume (12.3% of all non-marketable limit orders) submitted across all stock-days.

[8] Following BPV (see their Table 5), we use these variables to model the order exposure decision of each of our trader types. Results are reported in Internet Appendix D. They show that HFTs, AATs and NATs all react similarly to stock characteristics, order attributes, and market conditions.

[9] This method accounts for possible cross-correlations in the individual stock regressions. Assuming that the pairwise residual correlations are constant across stocks, Chordia et al. (2005) show that the usual standard error of the aggregate estimate is inflated by a factor $[1+(N\text{-}1)\rho]^{0.5}$, where $N$ is the number of



Electronic copy available at: https://ssrn.com/abstract=3074049

**[Table 4]**

Both buy and sell limit order execution models show consistent results. The $HFT_{ij}$ dummy has a negative and significant coefficient, indicating that HFTs cancel more orders before execution compared to NATs. The coefficient for the $HLO_{ij}$ dummy is negative (although only significant for buy orders) and explained by the fact that hidden orders lose time priority. The coefficient of interest is the dummy on $HLOHFT_{ij}$, which shows the execution probability of a hidden order placed by HFTs. HLOs placed by HFTs have a positive and significant coefficient for both buy (2.58) and sell (1.73) orders. In spite of the loss of time priority, HFTs' HLOs have higher execution probability, indicating that HFTs use HLOs effectively. In Panel B we show the unconditional execution probability of HLOs (and DLOs) for HFTs versus other traders. Compared to AATs and NATs, HFTs have higher rate of execution of their hidden orders submitted beyond the best quotes. This suggests that HFTs may be using non-aggressive HLOs in anticipation of volatility peaks or short-term order imbalances, a conjecture that we test later in this paper.

To complement the previous analysis on execution probability, we also examine the time to full execution of hidden orders placed by HFTs using survival analysis. Survival analysis can accommodate an important feature of limit order execution times: censored observations. If an order is cancelled 30 minutes after submission, then apparently it provides little information about the execution time, but the fact that it survived for 30 minutes is useful information. Such information contained in non-executed orders is used in survival analysis. We model the determinants of the execution of buy and sell limit orders separately and report the cross-sectional average estimates of the variables. The *t*-statistic for testing the significance of each variable is computed using the Chordia et al. (2005) method. We estimate the following model:

$$\begin{aligned} TIME_{ij} = {} & \alpha_0 + \alpha_1 DistMidQ_{ij} + \alpha_2 LastBuy_{ij} + \alpha_3 DepthSame_{ij} + \alpha_4 DepthSame^2{}_{ij} + \\ & + \alpha_5 DepthOpp_{ij} + \alpha_6 OrdSize_{ij} + \alpha_7 TrdFreq_{ij} + \alpha_8 RelTrdFreq_{ij} + \\ & + \alpha_9 HLO_{ij} + \alpha_{10} HFT_{ij} + \alpha_{11} AAT_{ij} + \alpha_{12} HLOHFT_{ij} + \alpha_{13} HLOAAT_{ij} + e_{ij} \end{aligned} \qquad [2]$$

where $TIME_{ij}$ is the time to full execution of the $j^{th}$ order, or the time survived in the book for a cancelled or expired order, with a positive censorship dummy.

---

stocks and $\rho$ is the common cross correlations. Since order arrival times vary across stocks, the regression residuals are not synched in time. To address this, we measure the average residual for each stock over 15-minute periods, and estimate $\rho$ as the average of 580 pairs of cross-correlation.



Electronic copy available at: https://ssrn.com/abstract=3074049

The model covariates are the same as in the previous analysis, and control for stock, order book, and market conditions, as well as the order placement strategy of the other trader types. As in the previous analysis, we exclude market and marketable limit orders and fleeting orders. The econometric specifications follow BPV and Lo et al. (2002) and model an accelerated failure time specification of limit order execution times under the generalized gamma distribution. The model is estimated on a stock-by-stock basis, and we report aggregated coefficients and significance levels. Results are in Table 5.

**[Table 5]**

Time to execution of both buy and sell limit orders show consistent results. As in BPV, NATs' HLOs take longer to fully execute (the coefficient of the HLO dummy is positive and statistically significant for both buy and sell orders). The coefficient of interest ($HLOHFT_{ij}$) which captures the relative time to full execution of an HLO placed by HFTs, has a negative and significant coefficient for both buy (-3.61) and sell (-2.76) orders, indicating that HFTs' HLOs take shorter time to fully execute compared to NATs, and AATs (the $HLOAAT_{ij}$ dummy is also negative but about half the magnitude compared to HFTs).

Together, the results in Tables 4 and 5 document that HFTs efficiently place their HLOs such that their time to execution is lower and their execution probability is higher. But at what cost? We next examine the costs HFTs face in their hidden order execution. To compute execution costs, it is important to note that iceberg orders are single (or parent) orders that are broken up into a sequence of smaller (child) orders. As the parent orders are executed, they are recorded in the data as multiple smaller transactions in a correlated sequence. However, as Perold (1988) pointed out, the cost incurred by the trader is not a function of a single transaction but rather the entire sequence of child orders. To accommodate this order splitting in cost computation, Perold (1988) introduced the implementation shortfall (ISF) metric, that measures transaction cost for the parent order. ISF compares the value of a paper portfolio with no transaction costs to the real portfolio obtained by actual trading. This method has been used in empirical work by Keim and Madhavan (1997), BPV, and Engle, Ferstenberg, and Russell (2012), among others. We use the ISF approach to evaluate the transaction costs of HLOs vis-à-vis DLOs for different trader types.



Electronic copy available at: https://ssrn.com/abstract=3074049

ISF for stock a given order *j* of a given stock *i* ($ISF_{ij}$) is the sum of the effective cost of execution or price impact ($PRI_{ij}$) and the opportunity costs of non-execution ($OPC_{ij}$). For a buy order

$$ISF_{ij} = PRI_{ij} + OPC_{ij} = \kappa_{ij} s_{ij} \left( \overline{p}_{ij} - q_0^i \right) + \left( 1 - \kappa_{ij} \right) s_{ij} \left( q_c^i - q_0^i \right), \qquad [3]$$

where the $PRI_{ij}$ component is the difference between the average execution price ( $\overline{p}_{ij}$ ) and the mid-quote at the time of order submission ( $q_0^i$ ), multiplied by the amount of shares executed ( $\kappa_{ij} s_{ij}$ ), where $s_{ij}$ is the order size (in shares) and $\kappa_{ij}$ is the fill rate of the order. The $OPC_{ij}$ for a buy order is the difference between the closing price on the day of order submission ( $q_c^i$ ) and $q_0^i$, multiplied by the unexecuted part of the order $\left( 1 - \kappa_{ij} \right) s_{ij}$ . Metrics for sell orders are analogously computed but conveniently signed.

Results are based on non-marketable limit orders. Revisions of standing limit orders are common in our data. We treat revisions of non-executed orders as the same order. In such cases, the *ISF* is computed using *s* as the order size after the last revision. Revisions of partially-executed orders are treated as new submissions. After computing the $ISF_{ij}$, $PRI_{ij}$, and $OPC_{ij}$ for each order *j*, we regress them on order attributes, market conditions during the 30 minutes prior to order submission, and trader-type dummies. Specifically, we estimate model [4] below using OLS regression on a stock-by-stock basis, with White-robust standard errors,

$$\begin{aligned} Y_{ij} = \alpha &+ \beta_A Aggr_{ij} + \beta_S OrdSize_{ij} + \beta_B Buy_{ij} + \beta_I HLO_{ij} + \\ &+ \beta_{HFT} HFT_{ij} + \beta_{AAT} AAT_{ij} + \beta_{IH} HLOHFT_{ij} + \beta_{IA} HLOAAT_{ij} + \\ &+ \beta_T TrdFreq_{ij} + \beta_T Volat_{ij} + e_{ij} \end{aligned} \qquad [4]$$

where $Y_{ij}$ is either $ISF_{ij}$, $PRI_{ij}$, or $OPC_{ij}$. The variable of interest is the $HLOHFT_{ij}$ dummy that captures the shortfall measure for HLOs placed by HFTs. We control for order attributes (order aggressiveness, size, direction), order types (HLO vs. DLOs), and trader categories (HFT, AAT, and NATs). See Internet Appendix C for variable definitions.

We report median estimated coefficients across stocks, the percentage of statistically significant coefficients, and the percentage of significant and positive coefficients. For the execution cost component ($PRI_{ij}$) we provide results conditional on partial execution (fill rate > 0%); for the opportunity costs component ($OPC_{ij}$), we provide results



Electronic copy available at: https://ssrn.com/abstract=3074049

conditional on non-full execution (fill rate < 100%). Note that a fully executed order has zero opportunity cost, and a completely non-executed order has zero execution cost. The results of the ISF analysis are reported in Table 6.

**[Table 6]**

In Panel A, we show the total $ISF_{ij}$ regression results. The estimated coefficient of $HLOHFT_{ij}$ is negative, significant for 39.29% of the sample, and positive only for 10.71% of these stocks. Thus, for a majority of our sample, the ISF of HFTs is either the same, or lower, than the ISF for NATs. This indicates that in spite of incurring higher execution costs, HFTs manage to substantially reduce their opportunity costs of non-trading, which offsets, and in many cases even outweighs, the total costs borne by traders with lower execution costs.

To probe how HFTs achieve reduced shortfall for their HLOs, we disaggregate the metric into its two components – effective costs of execution (Panel B) and opportunity cost of non-execution (Panel C). By construction, the effective costs component of any non-marketable buy (sell) limit order is negative (see the estimated negative intercept in Panel B), as they can only execute at a price below (above) the quote midpoint. For the same reason, the more aggressively priced the limit order, the higher (less negative) the effective costs component should be (see the positive coefficient of *Aggr* in Panel B). In Table 3, we saw that HFTs place their HLOs more aggressively than other traders. For all fill rates (all orders submitted), HFTs' effective costs of execution are higher than those of the NATs, both for DLOs and HLOs. The variable of interest, the HLOHFT dummy, has an average coefficient of 0.0126 and is significantly positive for a majority of stocks (64.29%). If we consider only those orders with fill rates greater than zero, that is orders that were at least partly executed, both the average HLOHFT estimated coefficient and the percentage of significantly positive coefficients increase. Our results therefore indicate that HFTs face higher effective costs than other traders when their HLOs are executed.

In contrast, in Panel C of Table 6, we find that the opportunity cost of non-execution for the HLOs of HFTs (HLOHFT dummy) is negative (average coefficient of -0.0714), and this result is stronger for fill rates under 100%. Thus, although HFTs' HLOs are executed at less favorable prices (larger effective costs), they experience less adverse



Electronic copy available at: https://ssrn.com/abstract=3074049

price movements in case of non-execution.[10] In sum, this latter (Panel C) effect dominates the former (Panel B), leading to an overall lower ISF (Panel A) result.

Overall, our findings in Tables 4, 5, and 6 indicate that HFTs manage the cost-benefit trade off involved in the order exposure decision better than AATs and NATs, resulting in more efficient use of HLOs.

## 6. The "why" question: Testing extant theory

In this section, we address the "why HFTs use HLOs" question by evaluating whether the free-option theory and/or the information-revelation theory can explain HFTs' order exposure decision.

### 6.1. The free option theory

To test the free-option theory of hidden order usage, we examine the order size distribution of HLOs (and DLOs) by trader type. Theory posits that traders who want to trade large positions hide their trading interest in order to limit the option value of their limit orders. Empirically BPV, among others, find this to be true in the pre-HFT era. We define trade size categories in total shares for both displayed and hidden orders and use the two-sample Kolmogorov-Smirnov (Massey, 1951) test to compare the order size distributions of HLOs and DLOs submitted by the different trader categories. Table 7 shows the hidden and displayed order sizes placed by HFTs, AATs, and NATs for large cap (Panel A), mid cap (Panel B) and small cap (Panel C) firms.

**[Table 7]**

In large cap firms (Panel A), the entry 76.28% under HFTs for HLOs indicates that 76.28% of HFT's HLOs are placed in the under-50-shares size category. By comparison, HFTs place only 5.11% of their displayed orders in this smallest share-size category and instead use larger share sizes when they expose their trading interest. Looking across the same row, we find that the pattern reverses for the NATs. These traders place more (65.99%) of their DLOs and less of their HLOs (29.13%) in this smallest size-category. Looking down each column, we find that the largest proportion of HFT's HLOs are in the smallest size category and this declines steeply as we move

[10] By definition informed traders have private signals about posterior changes in prices, so their limit orders should have higher opportunity costs of non-trading. Thus one interpretation of our finding is that HFTs' HLOs convey less information or, at least, they are less likely to be information-motivated. We provide formal evidence of this issue in the next section.



Electronic copy available at: https://ssrn.com/abstract=3074049

up to larger share brackets, with the largest (over 2500 shares) category receiving only 0.05% of the total HFTs' HLOs. DLOs, on the other hand, are more concentrated around the middle three share size categories (100-200, 200-500, and 500-1000 shares). NATs, by contrast, show a similar concentration around the middle order-size categories, but for their HLOs instead.

In mid-cap (Panel B) and small cap (Panel C) firms, HFTs place the majority of their HLOs – 98.72% and 83.96% respectively – in the smallest (under 50) share size category. The corresponding numbers for NATs – 31.53% and 22.77% – show that they do not hide as much of their orders in small share sizes. So, in the use of order sizes, we find a stark contrast between HFTs and NATs. While NAT's order size choice for hiding is consistent with previous literature, HFTs behave in quite the opposite way.

In each panel of Table 7, we also report the average size of HLOs and DLOs. In large stocks, for example, we find that while NATs place large order sizes (1139.59) as HLOs while they display their smaller (309.27) sized orders; HFTs do the opposite, they use large DLOs (1150.50) and comparatively smaller (459.58) HLOs.

These small-sized HLOs placed by HFTs bear out O'Hara's (2015) prescient summing up of the relationship between HFTs, small trades, and the ability to conceal trading interest that "small trade sizes reflect the influence of HFTs because [these] "silicon traders" can spot (and exploit) human traders by their tendency to trade in round numbers, [and] all trading is converging to ever smaller sizes and is being hidden whenever possible." The Nasdaq data also verify similar smaller order size usage by HFTs compared to non-HFTs (Table IB II).

Figure 2 plots the estimated cross-sectional daily average probabilities of HLO submission by HFTs, AATs, and NATs, conditional on order size and aggressiveness, for the large cap stocks. Figure 2.a shows that the HFTs' likelihood of hiding an order decreases with order size, which again is at odds with the free-option theory. While HFTs have a higher probability of placing small sized HLOs at all distances from the best quotes (at), they have the highest likelihood of placing such orders at the best quotes, followed by near the best quotes. These patterns reverse for both AATs (Figure 2.b) and NATs (Figure 2.c). Similar findings obtain in the Nasdaq data. For non-HFTs, the likelihood of hiding increases with order size. Fig IB 1 (in the Internet Appendix B) shows that HFTs place more HLOs in the smaller order size categories while the non-HFTs place more hidden volume using larger order sizes. Thus, evidence from both the



Electronic copy available at: https://ssrn.com/abstract=3074049

NSE and Nasdaq markets confirms that the free-option theory cannot explain HFTs' order exposure decision.

**[Figure 2]**

### 6.2. The information-revelation theory

To test the information-revelation theory of hidden order usage, we next examine the information content of HLOs vis-à-vis DLOs for different types of traders using three different metrics: the average permanent price impact based on Hasbrouck (1991a), the contribution to the order-flow-related component of the efficient price volatility based on Hasbrouck (1991b), and the share of price discovery based on Hasbrouck (1995). In these tests we use NSE data only since the Nasdaq data with HFT flags does not provide order by order records (but instead provides LOB snapshots).

#### 6.2.1. The permanent price impact approach

The evidence from non-high-frequency markets suggests that HLOs are generally uninformed (BPV, Pardo and Pascual, 2012). We next turn to the information content of HLOs placed by HFTs and compare that to the HLOs of AATs and NATs. We first calculate the permanent price impact of different types of orders placed by different types of traders. Unlike the multi-market settings in, for example, Huang (2002) and Barclay, Hendershott, and McCormick (2003), where price impact computations are affected by difficulties in trade-quote alignment, the fact that the NSE handles over 80% of the equity volume in Indian markets provide us the advantage of a consolidated market.

We estimate an extended version of the Structural Vector Autoregressive (VAR) model in Hasbrouck (1991a). The model is defined in event time ($t$), where an event may be a non-marketable limit order submission, cancellation of a standing limit order, or a trade (market or marketable limit order submission). Revisions that improve (degrade) prices or increase (decrease) quoted depth are treated as limit order submissions (cancellations). We distinguish between HFTs, AATs, and NATs, and for each trader type we consider two types of orders – HLOs and DLOs. As a result of these partitions, the VAR model has 13 equations: one for the quote midpoint return and 12 for order-flow related variables. As is usual in applications of Hasbrouck's method, we impose contemporaneous causality from the order flow to the quote midpoint revisions.



Electronic copy available at: https://ssrn.com/abstract=3074049

The optimal number of lags is determined using the Schwarz' Bayesian Information Criterion for each stock-day. We exclude stock-days that have less than 20 occurrences of any particular event, where the events are trades, HLOs, DLOs, or cancellations. The trade variable takes the value +1 (-1) for buyer- (seller-) initiated trades. Displayed and hidden submissions as well as cancellations that happen on the ask (bid) side of the LOB take the value (-1) +1. We reset the trading process at the end of each day, resetting all lagged values to zero. The model is estimated in event time, so contemporaneous correlation is negligible. Nonetheless, we compute the impulse-response functions (IRFs) such that any correlation is taken into account.[11]

In Panel A of Table 8 we report the IRFs obtained from the estimation process described above. The accumulated IRFs measure the average permanent price impact of an innovation to each trader type (in the columns) and type of event (along the rows) combination, computed as continuously compounded returns and presented in basis points. Estimates are cross-sectional stock-daily averages. Statistical significance is computed using standard errors clustered by stock and day (Thompson, 2011).

**[Table 8]**

As expected, trades have the largest estimated average price impact for all categories. Among the three trader types, HFT trades have the largest impact (1.2271), and this is significantly different from both NATs and AATs (boldfaced coefficients). Of greater interest to us, however, is the IRF of HLOs placed by HFTs, compared to the HLOs of the other two trader types. Here we find that HFTs' HLOs have a significant positive long-term price impact (coefficient of 0.1913 in Panel A) which is not significantly different from either AATs' (0.2401) or NATs' (0.2170).

Results in Panel A suggest that HLOs are more informative than DLOs, but this is true for all trader types. Notice, however, that these results do not control for the aggressiveness of limit orders. Earlier tests show that HFTs use more aggressive HLOs and their likelihood of hiding orders increases with order aggressiveness. To test whether the Panel A results are affected by order aggressiveness, in Panel B we show the IRFs for each trader type after controlling for their limit order aggressiveness.[12] We classify as aggressive (non-aggressive) any limit order placed at or within (beyond) the

[11] For full methodological details, see the Internet Appendix D.
[12] We exclude HFTs' non-aggressive HLOs from this analysis. There are few stock-days with enough occurrences of this type of event to allow us to include an extra equation in the VAR model.



Electronic copy available at: https://ssrn.com/abstract=3074049

prevailing best quotes. We find that the HLOs placed by HFTs have a positive but insignificant permanent price impact once we control for aggressiveness. Indeed, only their aggressive DLOs turn out to have a significant permanent price impact. In contrast, the estimated average price impacts of the HLOs placed by both AATs and NATs are positive and statistically significant even after including the controls. Our findings therefore suggest that the average HFTs' HLO has no significant information content.

### 6.2.2. The efficient price variance decomposition approach

The IRFs in Panels A and B of Table 8 provide a measurement of the average informativeness of HLOs placed by different trader types. But how important are these orders in the price formation process? If different trader types do not trade as frequently in each stock, then the IRFs will not be a good indicator of the information conveyed by each trader type's HLOs in the aggregate price formation process. To address this, we obtain an estimate of the relative contribution of HFTs', AATs', and NATs' trades and orders to the order-flow (OF)-related component of the efficient price.

For each stock-day, we estimate the efficient (or long-run) variance using the Vector Moving Average representation of the same 13-equation structural VAR model for quote midpoint changes and OF described before. Following the approach originally proposed by Hasbrouck (1991b), we split the efficient variance estimate into its OF-related and OF-unrelated components (see the Internet Appendix D for details). In Panel C of Table 8, we report the cross-sectional average estimated relative contribution of each type of event by each trader type (HFT, AAT, and NAT) to the OF-related efficient variance. Standard errors are clustered by both stock and day (Thompson, 2011). An event may be a non-marketable DLO submission, a non-marketable HLO submission, a cancellation of a standing limit order, or a trade (i.e., a market or marketable limit order submission). The magnitudes in Panel C sum up to 100%.

Overall, trades explain 67.05% of the OF-related price variance and DLOs explain another 25.95%. HLOs explain less than 8%. Looking across the columns, if we focus on all orders (last row of Panel C), it is clear that HFTs' orders (25.02%) contribute less than both AATs' (34.54%) and NATs' (40.44%) orders. Finally, the HLOs of HFT contribute the smallest (0.46%) to the OF-related efficient price variation. These results show that HFTs' HLOs convey less information into prices when compared with either their own DLOs, or with the DLOs and HLOs of AATs and NATs. The boldfaced



Electronic copy available at: https://ssrn.com/abstract=3074049

coefficients show that these differences between HFTs and other trader types are statistically significant.

### 6.2.3. The information share approach

Our third and final verification of the information-revelation theory is based on the information share (IS) approach of Hasbrouck (1995). Although much of the literature including Hasbrouck (1995) use this set-up in multi-market settings, it has also been used to assess IS across different trader/order categories (e.g., Hendershott and Riordan, 2013; Brogaard et al., 2019). In our particular application, we consider our six trader-type (HFTs, AATs, and NATs) order-type (DLOs and HLOs) combinations. For each, we collect the best ask and bid quotes at the end of each second and compute the quote midpoint. We assume that all the collected quotes share a common long-term component (the efficient price). The IS attributable to each trader-type order-type combination is the relative contribution of their innovations to the volatility of the common component. The ISs are estimated for each stock-day, but, as in previous tests, we report average IS across stock-days. Statistical significance is assessed using double-clustered standard errors.[13] Table 9 presents the results.

**[Table 9]**

HFTs' average IS is 30.85% for their DLOs and 6.13% for their HLOs. Both AATs and NATs have greater ISs for their HLOs, at 7.62% and 11.87% respectively. Tests of statistical significance show that these differences – both between HFTs and AATs and between HFTs and NATs – are significant.

Overall, the results in Tables 8 and 9 indicate that the information conveyed by the HLOs placed by HFTs is less than that of their DLOs, or the HLOs (and DLOs) of the other two trader types. Therefore, results do not support that the information-revelation can explain HFTs' order exposure decision.

## 7. The "why" question: Testing two conjectures

Since neither the free-option theory nor the information-revelation theory provides a framework to understand HFTs' order exposure decision, we call for new theory. To provide groundwork for such theory, we offer and empirically validate two conjectures about HFTs' motives for hiding orders. These conjectures are based on a survey of the

[13] For econometric details, see Internet Appendix D.



Electronic copy available at: https://ssrn.com/abstract=3074049

recent HFT literature (Hasbrouck, 2018; Hirschey, 2018) and not meant to be an exhaustive list of all possibilities.

### 7.1 Undercutting

Our first conjecture is that HFTs may use HLOs to undercut standing orders without being detected. HFTs with their super-fast computers are in a position to anticipate order flow (Angel and McCabe, 2013) and trade ahead of other investors' orders (Hirschey, 2018; Korajczyk and Murphy, 2019). Do they use HLOs to this end? If they do, are they more likely to undercut other traders' orders using HLOs than DLOs?

To address these questions, we consider all HLO and DLO submissions (including revisions that increase order size) for each trader-type order-type combination. We define an undercutting order as a limit order that (a) is placed immediately after another submission on the same side of the market, (b) comes in under 10 milliseconds of the previous order, and (c) improves the price of the previous one. In Panel A of Table 10, we report the percentage of HLOs and DLOs (per trader type) classified as undercutting orders. We present results using undercutting orders restricted to the five best quotes; however, our conclusions remain unchanged if we consider only the best quotes.

**[Table 10]**

Of the three trader types, HFTs use the highest proportion (5.0208% or 5.6019%) of HLOs to undercut orders within five ticks of the standing best quotes, both at a lower level of stock activity (at least 20 orders of each type – hidden or displayed – by each trader type – HFT, AAT, and NAT per stock-day) or at a higher level of stock activity (at least 50 orders, per order-type and trader-type, as defined above). Not surprisingly, they also use DLOs for undercutting, 3.009% or 2.6037% depending on the level of activity. Expectedly, NATs show the least amount of such undercutting activity, both for HLOs and DLOs.

This evidence, while illustrative, does not take account of market conditions. From BPV and our earlier regression results, we know that order exposure is affected by both stock and market attributes. Thus, we next estimate the logit regression in equation [5] to examine whether the observed higher rates of undercutting by HFTs' HLOs remain after controlling for market conditions and the state of the LOB.



Electronic copy available at: https://ssrn.com/abstract=3074049

$$
\begin{aligned}
U_{ij} = \beta_0 + \beta_1 DispSizeUnd_{ij} + \beta_2 AggrUnd_{ij} + \beta_3 HFT_{ij} + \beta_4 AAT_{ij} + \\
+\beta_5 HLOHFT_{ij} + \beta_6 HLOAAT_{ij} + \beta_7 HLONAT_{ij} + \beta_8 HidVol_{ij} + \\
+\beta_9 RSpr_{ij} + \beta_9 DepthSame_{ij} + \beta_{10} DepthOpp_{ij} + \beta_{11} Volat_{ij} + e_{ij}
\end{aligned} \quad [5]
$$

The dependent variable in this model is a dummy that takes the value of one if an order is an undercutting order as defined earlier, zero otherwise. The first two control variables describe the characteristics of the undercut order. First we consider the displayed size of the undercut order ($DispSizeUnd_{ij}$). We expect that when the undercut order has a larger displayed size, HFTs are more likely to jump ahead of it. Second we consider the aggressiveness of the undercut order ($AggrUnd_{ij}$). Aggressiveness is inversely captured by the number of ticks away from the best quote on the same side. The further the undercut order is from the best quotes, in other words less aggressive, the less likely it is to be undercut. Thus we expect a negative relationship between the aggressiveness of the undercut order and its chance of being undercut. We include the trader types $HFT_{ij}$ and $AAT_{ij}$ (NAT is captured in the intercept) and the interaction of trader categories with the $HLO_{ij}$ order type, plus relative spread ($RSpr_{ij}$), depth on same ($DepthSame_{ij}$) and opposite ($DepthOpp_{ij}$) sides, and volatility ($Volat_{ij}$), all as defined earlier. Finally, we include a variable that gauges the possibility of hidden order detection ($HidVol_{ij}$). This variable is a dummy that takes the value of one if the presence of hidden volume in the same side has been revealed, zero otherwise. Hidden volume is revealed at the time an undercutting order is placed if the quantity that has been traded at the prevailing best quote is greater than the displayed depth, which is only possible if there was additional (hidden) volume at the best quotes (e.g., Pardo and Pascual, 2012).

Panel B of Table 10 presents the results. The displayed size of the undercut order is positively related to the likelihood of undercutting, confirming that larger orders are more likely to be undercut. Likewise, when an order is closer to the top of the book (more aggressive), it is more likely to be undercut (shown by the negative and significant coefficient on $AggrUnd_{ij}$). $HidVol_{ij}$ is positive indicating that when traders can infer the presence of hidden volume at the best quotes, they are more likely to place orders to trade ahead of these HLOs. In fact, the odds ratio shows that this likelihood is 1.57 times (or 50.66% more) compared to the use of displayed orders by NATs (we use the DLOs of NATs as the reference group for all odds ratio calculation in this Panel).

The main variable of interest is $HLOHFT_{ij}$. The coefficient on this variable is 0.4149 and significant at the 1% level. Compare this to the negative coefficients on $HLOAAT_{ij}$



Electronic copy available at: https://ssrn.com/abstract=3074049

and $HLONAT_{ij}$. Clearly, HFTs use HLOs for undercutting, while the two other trader types are less likely to do the same. In fact, the odds ratio for HFTs is greater than 1 (1.5142) while for both other trader types it is lower than 1, indicating that while HFTs use HLOs to undercut the standing quotes at or near the top of the order book, the two other types are less likely to use HLOs for the same purpose.

### 7.2 Volatility anticipation

Another possible use of HLOs by HFTs could be to take advantage of volatility patterns. Previous results (see Panel B of Table 4) show that HFTs have the highest fill rate for non-aggressive HLOs, which suggests that they place their less aggressive orders in anticipation of short-term price fluctuations. When (transitory) volatility is high, limit orders will have a higher likelihood of execution and hiding such orders confers the added benefit of not revealing trading interest.

To test whether less aggressive HLOs placed by HFTs can predict peaks in short-term volatility, we regress 30- and 60-second post order submission volatility (quote mid-point volatility computed from LOB at one-second intervals) on the attributes of the submitted order. In particular, we estimate the following model,

$$\begin{aligned} RVolat_{ij} = {} & \beta_0 + \beta_1 LagRVolat_{ij} + \beta_2 HLO_{ij} + \beta_3 NonAggr + \beta_4 NonAggrHLO_{ij} + \\ & + \beta_5 HFT_{ij} + \beta_6 HLOHFT_{ij} + \beta_7 NonAggrHFT_{ij} + \beta_8 NonAggrHLOHFT_{ij} + \\ & + \beta_9 AAT_{ij} + \beta_9 HLOAAT_{ij} + \beta_{10} NonAggrAAT_{ij} + \beta_{11} NonAggrHLOAAT_{ij} + e_{ij} \end{aligned} \qquad [6]$$

We use dummy variables to distinguish between HLOs and DLOs, whether orders are from HFT, AAT or NATs, and whether orders are non-aggressive (*NonAggr*), i.e., submitted beyond the prevailing best quotes. All these control variables are as previously defined. Lagged volatility ($LagRVolat_{ij}$) is included since innovations in volatility are known to have serial correlation. The model is estimated at a pooled level for all stocks and orders, controlling for stock, day, and time of the day fixed effects. The coefficient of interest is on the triple interaction term $NonAggrHLOHFT_{ij}$, which captures the marginal effect of non-aggressive HLOs placed by HFTs. Our expectation, given the unusually high rates of these orders, is that they anticipate peaks in volatility. Results are summarized in Table 11.

**[Table 11]**



Electronic copy available at: https://ssrn.com/abstract=3074049

We find that non-aggressive HLOs of HFTs are submitted before high short-term volatility peaks. The coefficient on $NonAggrHLOHFT_{ij}$ is positive and significant, and is significantly greater than the coefficients on non-aggressive HLOs of non-HFTs, non-aggressive DLOs of HFT, and aggressive HLOs of HFTs. Results hold if we control for volatility persistence by including volatility before the order submission.

## 8. Conclusion

Regulators, market operators (exchanges), and investors all agree that transparency is a desirable property in financial markets. At the same time, research shows that there is such a thing as too much transparency.[14] Thus, all major exchanges allow traders to hide their trading interest by placing HLOs. To avoid a “corner solution” where everyone chooses to hide all trading intent, hidden volume faces a penalty in the form of losing time priority (i.e., HLOs are always ranked behind similarly priced DLOs, even if the HLO was submitted earlier).

Research on HLOs generally concludes that patient liquidity providers use the option to hide when they want to transact large quantities while avoiding picking off risks (De Winne and D’Hondt 2007, Buti and Rindi, 2013). These findings come from non-high frequency markets, or models that do not account for the use of HLOs by HFTs. Given that HFTs are the majority of traders and liquidity providers in many markets (the US, Japan, and Europe, for example) and an increasing fraction in many others (India, China, for example), whether and how they use the option to hide orders should be of interest.

In this paper we provide, to our knowledge, the first comprehensive account of hidden order use by HFTs. This study is made possible by our access to data from two sources: the NSE – the largest exchange in India that handles over 80% of the equity volume – which identifies in rich detail the types of traders as well as the order handling

[14] Bloomfield and O’Hara (1999) examine market transparency in a study tellingly titled “Market transparency: Who wins and who loses?” In this laboratory experiment they determine the effects of trade and quote disclosure on market efficiency, bid-ask spreads, and trader welfare. They find that although trade disclosure increases the informational efficiency of prices, it also increases opening bid-ask spreads by reducing market-makers' incentives to compete for order flow. As a result, trade disclosure benefits market makers at the expense of liquidity traders and informed traders. Additionally, they examine quote disclosure and find no discernible effects on market performance. Asquith, Au, Covert, and Pathak (2013) find that the introduction of the TRACE reporting system for bond markets helped some investors and dealers through a decline in price dispersion, while harming others through a reduction in trading activity.



Electronic copy available at: https://ssrn.com/abstract=3074049

system they use, and a non-public Nasdaq data sources that identifies HFTs and non-HFTs.

We find that HFTs make extensive use of HLOs. They do not appear to use HLOs to avoid picking-off risk but instead use small order sizes, placed nearer the top of the book using the non-display option. This pattern is different from the NATs, who hide large orders. We find that HFTs are more skilled at minimizing the implementation shortfall of their HLOs by reducing the opportunity costs of non-execution as well as improving the probability of execution.

We address the information content of HFTs' HLOs using three different measures to capture the information conveyed by such orders - the average permanent price impact, the contribution to the order-flow related component of the efficient price variance, and Hasbrouck's (1995) information share. All three metrics indicate that HLOs placed by HFTs have lower information content than their displayed orders, as well as the hidden orders of the other two trader types.

Collectively our evidence shows that HFTs' pattern of HLO use do not align with theoretical models of order exposure, and make a case for new theory. To that end, we offer and verify two conjectures to explain HFTs' HLO usage, noting that these are not meant to be an exhaustive list of all possible reasons HFTs may use HLOs. First, we find that HFTs use the non-display option to undercut standing orders at/near the best quotes, similar to the results in Hirschey (2018) where HFTs jump ahead of other investors' orders. Second, we show that HFTs are rather skilled at anticipating volatility and successfully place HLOs prior to volatility spikes, which increases the probability of their hidden order execution. By presenting novel and robust results on the use of HLOs by HFTs, we believe this study makes a useful contribution to the literature.



Electronic copy available at: https://ssrn.com/abstract=3074049

## References


Aitken, Michael J., Hank Berkman, and Derek Mak. 2001. The use of undisclosed limit orders on the Australian Stock Exchange, *Journal of Banking & Finance* 25 (8): 1589-1603.

Admati, Anat, and Paul Pfleiderer. 1988. A theory of intraday patterns: Volume and price variability. *The Review of Financial Studies* 1(1): 3-40.

Anand, Amber, Sugato Chakravarty, and Terrence Martell. 2005. Empirical evidence on the evolution of liquidity: Choice of market versus limit orders by informed and uninformed traders, *Journal of Financial Markets* 8(3): 288-308.

Angel, James J., and Douglas McCabe. 2013. Fairness in financial markets: The case of high frequency trading. *Journal of Business Ethics* 112(4): 585-595.

Asquith, Paul, Andrea Au, Thomas Covert, and Parag A. Pathak. 2013. The market for borrowing corporate bonds. *Journal of Financial Economics* 107(1): 155-182.

Barclay, Michael J., Terrence Hendershott, and D. Timothy McCormick. 2003. Competition among trading venues: Information and trading on electronic communications networks. *The Journal of Finance* 58(6): 2637-2665.

Baron, Matthew, Brogaard, Jonathan, Hagströmer, Björn, & Kirilenko, Andrei 2018. Risk and Return in High-Frequency Trading. *Journal of Financial and Quantitative Analysis*, forthcoming.

Bessembinder, Hendrik, Marios Panayides, and Kumar Venkataraman. 2009. Hidden liquidity: an analysis of order exposure strategies in electronic stock markets. *Journal of Financial Economics* 94(3), 361-383.

Bloomfield, Robert, and Maureen O'Hara. 1999. Market transparency: Who wins and who loses? *The Review of Financial Studies* 12(1): 5-35.

Bloomfield, Robert, Maureen O’Hara, and Gideon Saar. 2005. The “make or take” decision in an electronic market: Evidence on the evolution of liquidity. *Journal of Financial Economics* 75(1): 165-199.

Bloomfield, Robert, Maureen O'Hara, and Gideon Saar. 2015. Hidden liquidity: Some new light on dark trading. *Journal of Finance* 70(5), 2227-2274.

Boehmer, Ekkehart, Gideon Saar, Lei Yu. 2005. Lifting the veil: An analysis of pre-trade transparency at the NYSE. *Journal of Finance* 60(2): 783–813.

Boulatov, Alex. and Thomas. J. George. 2013. Hidden and displayed liquidity in securities markets with informed liquidity providers. *Review of Financial Studies* 26(8): 2096-2137.

Boussetta, Selma, Laurence D. Lescourret, and Sophie Moinas. 2017. The role of pre-opening mechanisms in fragmented markets. *SSRN Working Paper*.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan. 2014. High frequency trading and price discovery. *Review of Financial Studies* 27(8): 2267-2306.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan. 2019. Price discovery without trading: Evidence from limit orders. *Journal of Finance*, forthcoming.




Electronic copy available at: https://ssrn.com/abstract=3074049

Budish, Eric, Peter Cramton, and John Shim. 2015. The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics* 130(4): 1547-1621.

Buti, Sabrina, and Barbara Rindi. 2013. Undisclosed orders and optimal submission strategies in a limit order market. *Journal of Financial Economics* 109(3): 797-812.

Carrion, Allen. 2013. Very fast money: High-frequency trading on the NASDAQ. *Journal of Financial Markets* 16(4): 680-711.

Chakrabarty, Bidisha, Pamela Moulton, and Frank Wang. 2019. Attention effects in a high-frequency world. *SSRN Working Paper*.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam. 2005. Evidence on the speed of convergence to market efficiency. *Journal of Financial Economics* 76(2): 271-292.

Chordia, Tarun, T. Clifton Green, and Badrinath Kottimukkalur. 2018. Rent seeking by low latency traders: evidence from trading on macroeconomic announcements. *Review of Financial Studies* 31(12): 4650-4687.

Copeland, Thomas E., and Dan Galai. 1983. Information effects on the bid-ask spread. *The Journal of Finance* 38(5): 1457-1469.

De Winne, Rudy, and Catherine D'Hondt. 2007. Hide-and-seek in the market: placing and Detecting hidden orders. *Review of Finance* 11(4): 663-692.

Engle, Robert, Robert Ferstenberg, and Jeffrey Russell. 2012. Measuring and modeling execution cost and risk. *The Journal of Portfolio Management* 38(2): 14-28.

Fong, Kingsley Y.L., and Wai-Man Liu. 2010. Limit order revisions. *Journal of Banking & Finance* 34(8): 1873-1885.

Foucault, Thierry, Johan Hombert, and Ioanid Roşu. 2016. News trading and speed. *The Journal of Finance* 71(1): 335-382.

Glosten, Lawrence R., and Paul R. Milgrom. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14(1): 71-100.

Hagströmer, Björn, and Lars Nordén. 2013. The diversity of high-frequency traders. *Journal of Financial Markets* 16(4): 741-770.

Harris, Lawrence. 1997. Order exposure and parasitic traders. Unpublished working paper. University of Southern California, Los Angeles, CA.

Harris, Lawrence. 1998. Optimal dynamic order submission strategies in some stylized trading problems. *Financial Markets, Institutions, and Instruments* 7:1-76.

Hasbrouck, Joel. 1991a. Measuring the information content of stock trades, *Journal of Finance* 46(1): 179–207.

Hasbrouck, Joel. 1991b. The summary informativeness of stock trades: An econometric analysis, *Review of Financial Studies* 4(3): 571–595.

Hasbrouck, Joel. 1995. One security, many markets: Determining the contributions to price discovery, *Journal of Finance* 50(4): 1175–1199.

Hasbrouck, Joel. 2018. High-frequency quoting: Short-term volatility in bids and offers. *Journal of Financial and Quantitative Analysis* 53(2): 613-641.



Electronic copy available at: https://ssrn.com/abstract=3074049

Hasbrouck, Joel, and Gideon Saar. 2009. Technology and liquidity provision: The blurring of traditional definitions. *Journal of financial Markets* 12(2): 143-172.

Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld. 2011. Does algorithmic trading improve liquidity? *The Journal of Finance* 66(1): 1-33.

Hendershott, Terrence, and Ryan Riordan. 2013. Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis* 48(4): 1001-1024.

Hirschey, Nicholas. 2018. Do high-frequency traders anticipate buying and selling pressure? *SSRN Working Paper*.

Hoffmann, Peter. 2014. A dynamic limit order market with fast and slow traders. *Journal of Financial Economics* 113(1): 156-169.

Huang, Roger D. 2002. The quality of ECN and Nasdaq market maker quotes. *The Journal of Finance* 57(3): 1285-1319.

Kahraman, Bige, and Heather Tookes. 2017. Trader leverage and liquidity. *The Journal of Finance* 72(4): 1567-1610.

Kaniel, Ron, and Hong Liu. 2006. So what orders do informed traders use? *Journal of Business* 79(4): 1867-1913.

Keim, Donald B., and Ananth Madhavan. 1997. Transactions costs and investment style: an inter-exchange analysis of institutional equity trades. *Journal of Financial Economics* 46(3): 265-292.

Korajczyk, Robert A., and Dermot Murphy. 2019. High frequency market making to large institutional trades. *Review of Financial Studies* 32(3): 1034-1067.

Liu, Wai-Man. 2009. Monitoring and limit order submission risks. *Journal of Financial Markets* 12(1): 107-141.

Lo, Andrew W., A. Craig MacKinlay, and June Zhang. 2002. Econometric models of limit-order executions. *Journal of Financial Economics* 65(1): 31-71.

Massey Jr. Frank J. 1951. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association* 46(253): 68-78.

Moallemi, Ciamac, and Kai Yuan. 2017. A model for queue position valuation in a limit order book. *SSRN Working Paper*.

Moinas, Sophie. 2010. Hidden limit orders and liquidity in order driven markets. Unpublished working paper. TSE Working Paper Series 10-147.

O'Hara, Maureen. 2015. High frequency market microstructure. *Journal of Financial Economics* 116(2): 257-270.

O'Hara, Maureen, Chen Yao, and Mao Ye. 2014. What's not there: Odd lots and market data. *The Journal of Finance* 69(5): 2199-2236.

Pardo, Angel, and Roberto Pascual. 2012. On the hidden side of liquidity. *The European Journal of Finance* 18(10): 949-967.

Pascual, Roberto, and David Veredas. 2010. Does the open limit order book matter in explaining informational volatility? *Journal of Financial Econometrics* 8(1): 57–87.

Perold, Andre F. 1988. The implementation shortfall: Paper versus reality. *The Journal of Portfolio Management* 14(3): 4-9.



Electronic copy available at: https://ssrn.com/abstract=3074049

Ranaldo, Angelo. 2004. Order aggressiveness in limit order book markets. *Journal of Financial Markets* 7(1): 53-74.

Thompson, Samuel B. 2011. Simple formulas for standard errors that cluster by both firm and time. *Journal of financial Economics* 99(1): 1-10.

van Kervel, Vincent, and Albert J. Menkveld. 2019. High-frequency trading around large institutional orders. *The Journal of Finance*, forthcoming.

Weller, Brian M. 2018. Does algorithmic trading reduce information acquisition? *Review of Financial Studies* 31(6): 2184-2226.



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table I
## Sample descriptive statistics

This table provides daily cross-sectional average statistics for 100 stocks listed on the National Stock Exchange (NSE) of India. Cross-sectional averages are computed from daily averages per stock. The sample period is October to December 2013 (61 trading days). The sample comprises market-capitalization-based subsamples of 30 (largest), 40 (medium), and 30 (smallest) stocks. Market capitalization is daily average in billions of Rupees. Volume is in 10,000-share units, number of trades is in 100-trade units, depth is in 1000-share units, and Price is in Rupees. Daily volatility is {(maximum price/minimum price) –1}x100. The relative bid-ask spread is the ratio of the quoted spread to the quote midpoint, in basis points. The relative effective spread is two times the difference between the average trade price and the quote midpoint divided by the quote midpoint. Displayed (hidden) depth is the accumulated displayed (non-displayed) depth in the limit order book (LOB). MT is message traffic or the number of order messages (sum of submissions, cancellations, and revisions) in 1000-message units. We provide two proxies for HFT: the ratio of MT to trades (MT/Trd) and cancellations to trades (CAN/Trd). Share in MT denotes each trader type's share in message traffic. Liquidity metrics are generated from 1-minute snapshots of the LOB and averaged across observations. Statistical significance is evaluated using the non-parametric Wilcoxon rank-sum test. In Panel A, "***", "**", "*" on "Mid" ("Small") indicate statistically different from the "Large" ("Mid") subsample at the 1%, 5%, and 10% levels, respectively. In Panel B, significance under the AATs column tests for the difference between HFTs and AATs, and in the NATs column tests the differences between algorithmic traders (HFTs and AATs) and NATs.

Panel A: Sample statistics

| | | Market-capitalization-based subsamples | | |
|---|---|---|---|---|
| | Full sample | Large | Mid | Small |
| Market capitalization (billions) | 448.19 | 1464.64 | 20.2 *** | 2.39 *** |
| Volume ('0000) | 86.54 | 227.73 | 40.2 *** | 7.12 ** |
| Number of trades ('00) | 106.99 | 315.89 | 25.88 *** | 6.23 *** |
| Volatility | 42.36 | 32.96 | 44.37 *** | 49.08 |
| Relative bid-ask spread (bsp) | 44.24 | 8.7 | 42.87 *** | 81.61 *** |
| Displayed depth ('000) | 103.57 | 203.27 | 81.65 *** | 33.08 *** |
| Hidden depth ('000) | 25.62 | 50.58 | 19.5 *** | 8.81 |
| Price (Rupees) | 309.76 | 606.47 | 255.94 *** | 84.8 *** |

Panel B: Message traffic per trader type and subsample

| | | Trader types | | |
|---|---|---|---|---|
| Subsample | Variable | HFTs | AATs | NATs |
| Large | MT | 1191.19 | 139.26 * | 43.28 *** |
| | MT/Trd | 223.79 | 29.25 *** | 2.51 *** |
| | CAN/Trd | 10.32 | 1.99 *** | 0.27 *** |
| | Share in MT | 57.81 | 25.91 *** | 16.28 *** |
| Mid | MT | 6.37 | 11.19 * | 5.10 *** |
| | MT/Trd | 300.95 | 107.99 | 3.03 *** |
| | CAN/Trd | 30.96 | 1.77 * | 0.54 *** |
| | Share in MT | 16.72 | 43.16 *** | 40.11 *** |
| Small | MT | 0.77 | 3.51 *** | 1.31 *** |
| | MT/Trd | 94.75 | 146.18 *** | 3.62 *** |
| | CAN/Trd | 9.77 | 1.53 | 0.73 *** |
| | Share in MT | 5.97 | 57.23 *** | 36.80 *** |



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table II
## Use of HLOs

For a market-capitalization representative sample of 100 of NSE-listed stocks, this table provides cross-sectional average daily statistics on the use of hidden limit orders (HLOs) and displayed limit orders (DLOs) per trader type. We distinguish between high frequency traders (HFTs), agency algorithmic traders (AATs) and non-algorithmic traders (NATs) and provide statistics for subsamples of the largest (30), medium (40), and smallest (30) stocks in our sample. Our sample period is October to December 2013. In Panel A, we show the proportion of HLOs, both in the number of orders, and the accumulated volume, relative to all limit orders submitted. In Panel B, we provide each trader type's share of both HLOs and DLOs. Significant difference in medians between HFTs and AATs are shown beside AAT numbers and between all algorithmic traders (HFTs and AATs) and NATs are shown beside NAT numbers, using the non-parametric Wilcoxon rank-sum test. ***, **, * indicate statistically different at the 1%, 5%, and 10% levels respectively.

Panel A: Relative use of HLOs by type of trader and market capitalization subsample

| | | HFTs | | AATs | | NATs | |
|---|---|---|---|---|---|---|---|
| Variable | Subs. | Ord. | Vol. | Ord. | Vol. | Ord. | Vol. |
| % HLOs | Large | 10.38 | 9.83 | 24.31 *** | 32.47 *** | 8.86 * | 30.17 *** |
| | Mid | 36.00 | 34.42 | 15.97 | 26.40 | 14.48 ** | 34.03 |
| | Small | 15.84 | 15.23 | 3.36 *** | 7.77 *** | 13.38 *** | 32.42 *** |

Panel B: Market shares of HLOs and DLOs per trader type (%)

| Variable | Subs. | Orders | Volume | Orders | Volume | Orders | Volume |
|---|---|---|---|---|---|---|---|
| DLOs | Large | 34.67 | 55.84 | 21.59 * | 8.09 *** | 43.74 * | 36.07 *** |
| | Mid | 4.27 | 3.00 | 15.59 ** | 3.35 *** | 80.14 *** | 93.65 *** |
| | Small | 1.55 | 0.75 | 19.90 *** | 2.17 *** | 78.54 *** | 97.08 *** |
| HLOs | Large | 9.28 | 3.69 | 49.73 *** | 30.02 *** | 40.99 ** | 66.29 *** |
| | Mid | 18.90 | 8.14 | 13.45 | 6.34 ** | 67.65 *** | 85.52 *** |
| | Small | 5.80 | 2.49 | 4.36 *** | 1.57 *** | 89.84 *** | 95.95 *** |



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table III
## HLO placement in the order book

We examine the placement of hidden limit orders (HLOs) and displayed limit orders (DLOs), in Panels A and B respectively, both by the number of orders and the share volume. We build snapshots of the limit order book (LOB) at the time of each new order submission and group the LOB levels into four segments: (a) better than the standing quotes ("Better"), (b) at the best quotes ("At"), (c) from the best quotes up to 5 ticks away ("Near"), and (d) the rest ("Far"). The sample consists of 100 stocks listed on the NSE between October and December 2013 split into three market capitalization groups: largest (30), mid-sized (40), and smallest (30) stocks. We distinguish between high frequency traders (HFTs), agency algorithmic traders (AATs) and non-algorithmic traders (NATs). Each statistic reported is the time series mean of the daily proportion of orders at the four LOB level groups for all stocks taken together. We average ask and bid quotes. Statistical tests compare the medians of corresponding groups across Panels A and B, using the non-parametric Wilcoxon rank-sum test. (***, **, * indicate statistically different at the 1%, 5%, and 10% level respectively).

| | | Order placement | | | Volume placement | | |
|---|---|---|---|---|---|---|---|
| Subsample | Aggressiveness | HFTs | AATs | NAT | HFTs | AATs | NAT |
| Panel A: HLOs placement | | | | | | | |
| Large | Better | 25.11 *** | 10.81 *** | 12.05 *** | 26.64 *** | 7.55 *** | 6.80 *** |
| | At | 20.92 *** | 37.58 *** | 17.72 *** | 30.55 *** | 36.57 *** | 29.11 *** |
| | Near | 51.68 *** | 38.03 ** | 31.11 *** | 39.88 *** | 34.73 *** | 29.80 *** |
| | Far | 2.28 *** | 13.57 *** | 39.12 *** | 2.93 *** | 21.15 *** | 34.30 *** |
| Mid | Better | 70.14 *** | 43.57 *** | 20.27 *** | 72.01 *** | 13.29 | 10.57 *** |
| | At | 15.66 | 26.76 ** | 16.92 | 13.31 *** | 45.20 *** | 22.68 *** |
| | Near | 14.02 *** | 22.57 *** | 27.35 *** | 14.43 *** | 28.07 *** | 24.65 |
| | Far | 0.19 *** | 7.10 | 35.45 *** | 0.25 *** | 13.43 | 42.10 *** |
| Small | Better | 82.33 *** | 49.34 *** | 25.98 *** | 85.07 *** | 26.96 * | 16.43 *** |
| | At | 5.60 *** | 15.31 | 14.39 *** | 4.97 *** | 33.12 *** | 17.45 |
| | Near | 11.82 *** | 31.62 *** | 27.31 *** | 9.79 *** | 32.62 *** | 24.34 ** |
| | Far | 0.25 *** | 3.74 *** | 32.31 *** | 0.18 *** | 7.29 *** | 41.78 *** |
| Panel B: DLOs placement | | | | | | | |
| Large | Better | 0.47 | 2.43 | 4.32 | 0.08 | 0.64 | 2.40 |
| | At | 1.03 | 5.01 | 7.26 | 0.83 | 3.27 | 21.00 |
| | Near | 9.17 | 35.53 | 13.24 | 5.92 | 21.80 | 18.87 |
| | Far | 89.32 | 57.03 | 75.18 | 93.18 | 74.29 | 57.72 |
| Mid | Better | 36.17 | 23.06 | 11.56 | 3.28 | 13.71 | 7.26 |
| | At | 16.55 | 22.18 | 16.80 | 5.12 | 30.54 | 32.09 |
| | Near | 31.36 | 49.53 | 22.17 | 37.25 | 43.89 | 24.66 |
| | Far | 15.92 | 5.23 | 49.47 | 54.35 | 11.87 | 35.99 |
| Small | Better | 30.88 | 24.52 | 13.17 | 21.05 | 17.53 | 9.32 |
| | At | 14.77 | 13.24 | 12.23 | 15.24 | 17.47 | 18.59 |
| | Near | 37.23 | 55.37 | 20.82 | 40.17 | 46.78 | 22.77 |
| | Far | 17.12 | 6.87 | 53.78 | 23.54 | 18.22 | 49.32 |



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table IV
## Likelihood of order execution

We study the determinants of execution of hidden limit orders (HLOs) and displayed limit orders (DLOs) in the NSE. We distinguish between high frequency traders (HFTs), agency algorithmic traders (AATs) and non-algorithmic traders (NATs). To model order execution likelihood, in Panel A we use an ordered Logit model, where the dependent variable (EXEC) is an ordinal variable that takes three possible values: EXEC = 1 indicates that the limit order is cancelled before execution; EXEC = 2 indicates that the limit order is partially executed and then cancelled; EXEC = 3 indicates that the limit order is fully executed. The models are estimated on a stock-by-stock basis, and we report aggregated coefficients and significance levels based on Chordia, Roll, and Subrahmanyam (2005). In Panel B we show the unconditional execution likelihood for HLOs and DLOs placed at different levels of the limit order book. We consider three levels relative to the best quotes. The estimation sample consists of the 30 largest stocks (in which HFTs are reasonably active) from our main sample of 100 stocks listed on the NSE. The sample period is December 2013. In Panel A ***, **, * indicate significance at the 1%, 5% and 10% level, respectively. In Panel B ***, **, * indicate significantly different from HFTs at the 1%, 5% and 10% level, respectively.

Panel A: Likelihood of execution - Ordered probit model

| | Limit order to buy | | Limit order to sell | |
|---|---|---|---|---|
| Variable | Coef. | | Coef. | |
| Aggr | 273.7968 | *** | 159.7378 | *** |
| OrdSize | -2405.2770 | ** | -2055.9926 | ** |
| HLO | -0.4301 | ** | -0.2509 | |
| HFT | -2.2576 | *** | -2.2935 | *** |
| AAT | -1.6361 | *** | -1.4190 | *** |
| HLOHFT | 2.5816 | *** | 1.7313 | *** |
| HLOAAT | 1.3648 | *** | 0.9437 | *** |
| RSpr | 530.1330 | *** | 490.8146 | *** |
| DepthSame | -85.8532 | *** | -59.6091 | ** |
| DepthOpp | 62.5362 | *** | 72.1912 | *** |
| LOBImb | -0.1518 | *** | 0.1550 | *** |
| LastHalfHour | 0.2398 | *** | 0.2658 | *** |
| OI | -0.1464 | *** | 0.1131 | ** |
| TrdFreq | 1.1414 | ** | 1.4850 | ** |
| Mom | 7.7690 | | 3.9942 | |
| Volat | 4897.55 | | 6661.67 | |

Panel B: Likelihood of execution and order placement

| Placement/trader type | HLOs | | DLOs | |
|---|---|---|---|---|
| At or within the best quotes: | | | | |
| HFT | 79.10 | | 79.42 | |
| AAT | 86.82 | | 71.19 | ** |
| NAT | 85.34 | | 86.48 | |
| Within the 2nd and 5th best quotes: | | | | |
| HFT | 83.42 | | 48.40 | |
| AAT | 56.15 | *** | 32.61 | ** |
| NAT | 66.53 | *** | 73.12 | *** |
| Beyond the 5th best quote | | | | |
| HFT | 81.84 | | 6.98 | |
| AAT | 25.47 | *** | 24.75 | *** |
| NAT | 51.78 | *** | 50.51 | *** |



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table V
## Time to completion: Survival analysis

We study the determinants of the time to full execution of non-marketable limit orders at the NSE. We exclude market and marketable limit orders. We also drop fleeting orders (as defined by Hasbrouck and Saar, 2009). Revisions of non-executed orders are treated as the same order. Revisions of partially-executed orders are treated as new submissions. The table reports the estimated parameters of an econometric model of time-to-completion using survival analysis. We follow Bessembinder et al. (2009) and Lo, et al. (2002). The model describes an accelerated failure time specification of limit order execution times under the generalized gamma distribution. The model is estimated on a stock-by-stock basis, and we report aggregated coefficients and significance levels based on Chordia, Roll, and Subrahmanyam (2005). The estimation sample for this table consists of the 30 largest stocks (in which HFTs are reasonably active) from our main sample of 100 stocks listed on the NSE of India and the sample period is December 2013. ***, **, * indicate significance at the 1%, 5% and 10% level, respectively.

| Variable | Limit order to buy | Limit order to sell |
|---|---|---|
| Intercept | 16.7115 *** | 17.0854 *** |
| DistMidQ | 2.8183 ** | -1.6744 *** |
| LastBuy | 0.0762 * | -0.0918 |
| DepthSame | 227.3927 *** | 221.0423 ** |
| $\text{DepthSame}^2$ | -169.5214 ** | -151.7108 ** |
| DepthOpp | -196.9867 *** | -227.5512 *** |
| OrdSize | 47.1514 *** | 37.3681 ** |
| TrdFreq | -14.2400 ** | -10.3429 * |
| RelTrdFreq | -1.5036 *** | -1.4494 *** |
| HLO | 1.4503 *** | 1.1420 *** |
| HFT | 2.7756 *** | 2.4768 *** |
| AAT | 0.4430 | 0.0509 |
| HLOHFT | -3.6125 *** | -2.7638 *** |
| HLOAAT | -1.5064 *** | -1.1791 ** |



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table VI
## Implementation shortfall of HLOs

We present the effective costs of execution and the opportunity costs of non-execution costs of hidden limit orders (HLOs) and displayed limit orders (DLOs) in the NSE using the implementation shortfall (ISF) approach of Perold (1988). Execution cost for a buy order is the difference between the average execution price and the mid-quote at the time of order submission, multiplied by the amount of shares executed. The opportunity cost for a buy order is the difference between the closing price on the day the order is cancelled or expires and the quote midpoint at the time the order is submitted, multiplied by the unexecuted part of the order (in shares). Metrics for sell orders are analogously computed but conveniently signed. We regress each cost component on order attributes (order aggressiveness, total size, buyer order indicator, and HLO indicator), market conditions during the 30 minutes prior to order submission (trading frequency and realized volatility), and trader-type dummies. We estimate regressions for the whole ISF, but also for the execution cost component (PRI), and the opportunity costs component (OPC) separately. Models are estimated on a stock-by-stock basis. We report median estimated coefficients across stocks, the percentage of statistically significant coefficients, and the percentage of significant and positive coefficients. Note that a fully executed order has zero opportunity cost, and a fully cancelled order has zero execution cost. For the execution cost component we provide results conditional on partial execution (fill rate > 0%); for the opportunity costs component, we provide results conditional on non-full execution (fill rate < 100%). The estimation sample for this table consists of the 30 largest stocks (in which HFTs are reasonably active) from our main sample of 100 stocks listed on the NSE of India and the sample period is December 2013. We consider only non-marketable limit orders. Revisions of non-executed orders are treated as the same order. Revisions of partially-executed orders are treated as new submissions.



Electronic copy available at: https://ssrn.com/abstract=3074049

Panel A: Implementation shortfall (ISF)

| | All fill rates | |
|---|---|---|
| Variable | Coef. | %Signif.(pos.) |
| Intercept | 0.0638 | 92.86 (60.71) |
| Aggr | 1.2400 | 82.14 (46.43) |
| OrdSize | -78.4506 | 75.00 (14.29) |
| Buy | -0.1703 | 96.43 (28.57) |
| HLO | 0.0121 | 53.57 (39.29) |
| HFT | 0.0348 | 78.57 (50.00) |
| AAT | 0.0141 | 75.00 (50.00) |
| HLOHFT | -0.0445 | 39.29 (10.71) |
| HLOAAT | -0.0016 | 64.29 (35.71) |
| TrdFreq | 0.1042 | 64.29 (50.00) |
| Volat | 77.2356 | 60.71 (32.14) |

Panel B: Effective costs (PRI)

| | All fill rates | | Fill rate >0% | |
|---|---|---|---|---|
| Intercept | -0.0315 | 92.86 (0.00) | -0.0151 | 85.71 (0.00) |
| Aggr | 0.4152 | 89.29 (75.00) | 12.9157 | 92.86 (92.86) |
| OrdSize | -5.8493 | 82.14 (0.00) | -70.6004 | 82.14 (0.00) |
| Buy | 0.0035 | 89.29 (78.57) | 0.0021 | 78.57 (53.57) |
| HLO | -0.0139 | 89.29 (10.71) | -0.0128 | 71.43 (10.71) |
| HFT | 0.0297 | 89.29 (89.29) | 0.0081 | 67.86 (64.29) |
| AAT | 0.0338 | 92.86 (92.86) | 0.0267 | 92.86 (92.86) |
| HLOHFT | 0.0126 | 71.43 (64.29) | 0.0503 | 75.00 (71.43) |
| HLOAAT | 0.0134 | 92.86 (78.57) | 0.0184 | 85.71 (78.57) |
| TrdFreq | -0.0095 | 82.14 (7.14) | -0.0074 | 75.00 (14.29) |
| Volat | -164.6489 | 85.71 (10.71) | -68.6062 | 39.29 (0.00) |

**Table VI (Cont.)**
**Implementation shortfall of HLOs**

Panel C: Opportunity costs of non-execution (OPC)

| | All fill rates | | Fill rate <100% | |
|---|---|---|---|---|
| Intercept | 0.0799 | 92.86 (64.29) | 0.1433 | 92.86 (64.29) |
| Aggr | 0.0653 | 75.00 (35.71) | 0.0021 | 75.00 (42.86) |
| OrdSize | 0.6778 | 57.14 (25.00) | 4.9032 | 53.57 (25.00) |
| Buy | -0.1714 | 89.29 (28.57) | -0.2998 | 89.29 (28.57) |
| HLO | 0.0432 | 67.86 (53.57) | 0.1359 | 53.57 (39.29) |
| HFT | -0.0033 | 78.57 (32.14) | -0.0358 | 71.43 (21.43) |
| AAT | -0.0182 | 67.86 (14.29) | -0.0744 | 78.57 (14.29) |
| HLOHFT | -0.0714 | 46.43 (10.71) | -0.1022 | 50.00 (10.71) |
| HLOAAT | -0.0292 | 82.14 (32.14) | -0.0633 | 67.86 (25.00) |
| TrdFreq | 0.1159 | 71.43 (53.57) | 0.4274 | 67.86 (53.57) |
| Volat | 193.9898 | 57.14 (32.14) | 239.0775 | 67.86 (35.71) |



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table VII
## Order size

We provide cross-sectional average daily statistics on the empirical distribution of the size of hidden limit orders (HLOs) and displayed limit orders (DLOs) in the NSE. The sample consists of 100 stocks listed on the NSE between October and December 2013 that we split into three market capitalization groups: large caps (Panel A), mid-sized (Panel B), and small caps (Panel C), of sizes 30, 40, and 30 stocks, respectively. We distinguish between high frequency traders (HFTs), agency algorithmic traders (AATs) and non-algorithmic traders (NATs). The analysis is based on order-by-order data that we group according to the full (displayed plus non-displayed) order size. Trade size categories are defined in total (both displayed and hidden) shares. We use the two-sample Kolmogorov-Smirnov (Massey, 1951) test to compare the order size distributions of HLOs and DLOs submitted by the different trader categories. We provide the percentage of HLOs and DLOs in each order-size category per trader type. ***, **, * indicate statistically different than the corresponding HFTs' statistic at the 1%, 5% and 10% level, respectively.



Electronic copy available at: https://ssrn.com/abstract=3074049

| | HFTs | | AATs | | NATs | |
|---|---|---|---|---|---|---|
| **Panel A. Large caps** | | | | | | |
| Order size distrib. (%) | DLOs | HLOs | DLOs | HLOs | DLOs | HLOs |
| (0,50] | 5.11 | 76.28 | 60.17 | 55.23 | 65.99 | 29.13 |
| (50,75] | 0.79 | 10.15 | 10.91 | 8.25 | 1.52 | 2.69 |
| (75,100] | 1.19 | 0.55 | 4.18 | 6.25 | 11.14 | 11.19 |
| (100,200] | 22.01 | 2.24 | 11.42 | 12.11 | 6.36 | 11.98 |
| (200,500] | 46.53 | 7.91 | 10.40 | 11.33 | 9.03 | 22.26 |
| (500,1000] | 19.02 | 2.39 | 1.54 | 4.03 | 3.26 | 10.79 |
| (1000,2500] | 2.82 | 0.44 | 0.73 | 2.03 | 1.46 | 6.07 |
| >2500 | 2.53 | 0.05 | 0.65 | 0.77 | 1.24 | 5.89 |
| HFTs vs. AATs/NATs (p-value) | | | | 0.00 | | 0.00 |
| DLOs vs. HLOs (p-value) | | 0.00 | | 0.00 | | 0.00 |
| Average size (sh.) | 1150.50 | 459.58 | 345.36 *** | 880.65 * | 309.27 ** | 1139.59 *** |
| **Panel B. Mid-sized caps** | | | | | | |
| (0,50] | 62.98 | 98.72 | 71.60 | 63.84 | 51.81 | 31.53 |
| (50,75] | 3.19 | 1.03 | 9.48 | 6.67 | 1.86 | 1.69 |
| (75,100] | 7.19 | 0.24 | 8.60 | 4.61 | 14.21 | 13.44 |
| (100,200] | 8.23 | 0.01 | 4.74 | 8.79 | 9.87 | 10.01 |
| (200,500] | 6.09 | 0.00 | 4.56 | 8.82 | 12.86 | 19.83 |
| (500,1000] | 1.55 | 0.00 | 0.59 | 4.00 | 5.00 | 10.06 |
| (1000,2500] | 0.81 | 0.00 | 0.27 | 2.11 | 2.38 | 6.61 |
| >2500 | 9.96 | 0.00 | 0.15 | 1.17 | 2.01 | 6.84 |
| HFTs vs. AATs/NATs (p-value) | | | | 0.00 | | 0.00 |
| DLOs vs. HLOs (p-value) | | 0.00 | | 0.00 | | 0.00 |
| Average size (sh.) | 207.85 | 94.35 | 96.68 | 1342.57 *** | 396.23 *** | 1247.97 *** |
| **Panel C. Small caps** | | | | | | |
| (0,50] | 46.51 | 83.96 | 87.99 | 75.64 | 47.95 | 20.77 |
| (50,75] | 4.19 | 15.44 | 1.43 | 3.16 | 1.58 | 1.75 |
| (75,100] | 29.71 | 0.58 | 7.82 | 3.20 | 14.85 | 16.32 |
| (100,200] | 12.24 | 0.00 | 1.62 | 6.11 | 11.10 | 12.46 |
| (200,500] | 5.85 | 0.01 | 0.84 | 5.62 | 15.76 | 22.51 |
| (500,1000] | 1.07 | 0.00 | 0.20 | 3.02 | 5.23 | 12.22 |
| (1000,2500] | 0.41 | 0.00 | 0.08 | 2.16 | 2.15 | 6.96 |
| >2500 | 0.01 | 0.00 | 0.03 | 1.10 | 1.38 | 7.03 |
| HFTs vs. AATs/NATs (p-value) | | | | 0.00 | | 0.00 |
| DLOs vs. HLOs (p-value) | | 0.00 | | 0.00 | | 0.00 |
| Average size (sh.) | 127.81 | 99.06 | 49.52 | 740.69 *** | 319.05 *** | 1196.78 ** |



Electronic copy available at: https://ssrn.com/abstract=3074049

# Table VIII
## Impulse-response functions and order-flow related variance decomposition

In Panels A and B we provide stock-day average impulse response functions (IRF) from an extended VAR (Hasbrouck, 1991a). In Panel C we estimate the efficient variance using the Hasbrouck (1991b) approach and decompose the efficient variance into an order-flow-related component and an order-flow-unrelated component. For all models we use order level data for December 2013 on the 30 largest stocks in our representative sample of 100 NSE-listed stocks. The models are defined in event time (*t*), where an event may be a limit order submission, cancellation, or trade. Revisions that improve (degrade) prices or increase (decrease) quoted depth are treated as limit order submissions (cancellations). We distinguish between high frequency traders (HFTs), agency algorithmic traders (AATs), and non-algorithmic traders (NATs). We differentiate between hidden (HLOs) and displayed limit orders (DLOs). As a result of these partitions, the models have 13 equations: one for the quote midpoint return and 12 for order-flow related variables. The optimal number of lags is determined using the Schwarz' Bayesian Information Criterion. "Trade" variables are signed +1 (-1) for buyer- (seller-) initiated trades. "DLO", "HLO" or "Cancellation" variables that happen on the ask (bid) side of the LOB are signed (-1) +1. We assume the trading process restarts each day, resetting all lagged values to zero. Standard errors are clustered by both stock and day (Thompson, 2011). In Panel B, present the IRF tests controlling for order aggressiveness(a) versus non-aggressiveness (na). ***, **, * indicate significance at the 1%, 5% and 10% level, respectively. In Panels B and C, we boldface those coefficients for AATs and NATs that are significantly different from corresponding coefficients for HFTs.

Panel A: IRF - Continously-compound return (in basis points)

| Message | All traders | HFT | AAT | NAT |
|---|---|---|---|---|
| | | Trader type | | |
| Trades | | 1.2271 *** | **0.7259** *** | **0.8582** *** |
| | | (0.1382) | (0.1017) | (0.1474) |
| DLO | | 0.0816 ** | 0.0568 *** | **0.1640** *** |
| | | (0.0318) | (0.0099) | (0.0260) |
| HLO | | 0.1913 *** | 0.2401 *** | 0.2170 *** |
| | | (0.0536) | (0.0328) | (0.0308) |
| Cancellations | | 0.0793 *** | 0.0454 *** | **0.1233** *** |
| | | (0.0291) | (0.0117) | (0.0254) |

Panel B: IRF - controlling for aggressiveness

| Message | All traders | HFT | AAT | NAT |
|---|---|---|---|---|
| Trades | | 1.1591 *** | **0.7273** *** | **0.8583** *** |
| | | (0.1261) | (0.1039) | (0.1485) |
| DLOa | | 0.2512 *** | 0.2410 *** | **0.6221** *** |
| | | (0.0505) | (0.0288) | (0.0696) |
| DLOna | | 0.0111 * | -0.0014 | -0.0002 |
| | | (0.0064) | (0.0052) | (0.0039) |
| HLOa | | 0.1778 | **0.3523** *** | **0.4907** *** |
| | | (0.1132) | (0.0445) | (0.0543) |
| HLOna | | | -0.0351 | -0.0239 ** |
| | | | (0.0225) | (0.0116) |
| Cancellations | | 0.0623 *** | 0.0502 *** | **0.1163** *** |
| | | (0.0196) | (0.0100) | (0.0238) |

Panel C: OF-related efficient variance (OFEV) decomposition

| Message | All traders | HFT | AAT | NAT |
|---|---|---|---|---|
| Trades | 67.05 | 16.09 *** | **21.39** *** | **29.57** *** |
| | | (1.69) | (3.13) | (2.24) |
| Limit orders | 25.95 | 6.18 *** | **9.25** *** | **10.52** *** |
| | | (1.03) | (1.21) | (0.93) |
| HLOs | 7.84 | 0.46 ** | **5.68** *** | **1.69** *** |
| | | (0.18) | (0.87) | (0.14) |
| Cancellations | -0.84 | 2.29 *** | **-1.78** ** | **-1.34** *** |
| | | (0.72) | (0.73) | (0.25) |
| All orders | | 25.02 | 34.54 | 40.44 |



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table IX
## Information shares

The table reports the average stock-day information shares (IS) for different types of traders and orders in the NSE. Information shares are estimated using Hasbrouck (1995) approach. We report lower bound (minimum), upper bound (maximum), and average information shares for three types of traders: proprietary ATs (hereafter, HFTs), agency ATs (hereafter, AATs), and non-ATs (hereafter, NATs). Moreover, we distinguish between hidden limit orders (HLOs) and fully displayed limit orders (DLOs). On a one-second frequency, we obtain the best quotes for each trader type and order type. The price path of each trader type and order type pair is given by the quote midpoint prevailing at the end of each second. Using the IS approach, we decompose the variation in the unobserved common efficient price into individual components attributable to specific trader and order type. Our main purpose is to examine the fraction of price discovery attributable to HLOs and how much of it is attributable to HFTs' and ATs' orders. We use order level data for December 2013 on the 30 largest stocks in our representative sample of 100 NSE-listed stocks. ***, **, * next to a HFTs' or ATs' IS indicates that the IS statistic is significantly different from the corresponding NATs' IS statistic for the same order type.

| | | Information shares (%) | | |
|---|---|---|---|---|
| Trader type | Order | Min. | Max. | Avg. |
| HFTs | DLO | 15.87 | 45.83 | 30.85 |
| | HLO | 5.91 | 6.34 *** | 6.13 ** |
| AATs | DLO | 8.81 *** | 34.44 *** | 21.62 *** |
| | HLO | 5.00 | 10.25 *** | 7.62 ** |
| NATs | DLO | 16.22 | 47.62 | 31.92 |
| | HLO | 6.36 | 17.39 | 11.87 |



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table X
## Undercutting using HLOs

We present the proportions of hidden (and displayed) orders used for undercutting by the three trader categories (in Panel A) and use a logit regression model to study the likelihood of undercutting by these three trader types (in Panel B). We define an undercutting limit order as a limit order that (a) is placed immediately after another submission on the same side of the market, (b) comes in under 10 milliseconds of the previous order, and (c) improves the price of the previous one. We present results using undercutting orders restricted to the five best quotes. We divide the total number of undercutting orders of each type – hidden and displayed – placed by each trader type – HFT, AAT, and NAT – by all orders submitted of a given type by each trader type. We present those fractions in Panel A. In Panel B we present the coefficients and odds ratios of the logit regression where the dependent variable is a dummy that takes the value of 1 if the order is an undercutting order, 0 otherwise. The models are estimated on a stock-by-stock basis, and we report aggregated coefficients and t-statistics using the approach in Chordia, Roll, and Subrahmanyam (2005). The estimation sample for this table consists of the 30 largest stocks (in which HFTs are reasonably active) from our main sample of 100 stocks listed on the NSE of India and the sample period is December 2013. ***, **, * indicate significance at the 1%, 5%, and 10% level respectively.

Panel A: Descriptive statistics on undercutting (% of orders)

| First case: At least 20 orders per category and stock-day | | | |
|---|---|---|---|
| Order | TraderType | Bid side | Ask side |
| HLO | HFT | 5.0208 *** | 5.4092 *** |
| | AAT | 3.2303 *** | 3.4066 *** |
| | NAT | 0.8173 *** | 0.8078 *** |
| DLO | HFT | 3.0091 *** | 3.2427 *** |
| | AAT | 4.6264 *** | 5.0179 *** |
| | NAT | 1.1373 *** | 1.1707 *** |
| **Second case: At least 50 orders per category and stock-day** | | | |
| Order | TraderType | Bid side | Ask side |
| HLO | HFT | 5.6019 *** | 6.0651 *** |
| | AAT | 3.3964 *** | 3.4847 *** |
| | NAT | 0.8088 *** | 0.8025 *** |
| DLO | HFT | 2.6037 *** | 2.7307 *** |
| | AAT | 5.1687 *** | 5.5820 *** |
| | NAT | 1.0611 *** | 1.0792 *** |

Panel B: Logit model on undercutting

| Variable | Coef. | Odds ratio | CRS t-stat |
|---|---|---|---|
| DispSizeUnd | 0.0004 *** | 1.0004 | 10.03 |
| AggrUnd | -0.0744 *** | 0.9283 | -119.14 |
| HFT | 0.7620 *** | 2.1425 | 39.49 |
| AAT | 0.9856 *** | 2.6794 | 40.69 |
| HLOHFT | 0.4149 *** | 1.5142 | 7.67 |
| HLOAAT | -0.1902 | 0.8268 | -0.06 |
| HLONAT | -0.5556 *** | 0.5737 | -3.96 |
| HidVol | 0.4489 *** | 1.5666 | 66.72 |
| RSpr | 0.0300 *** | 1.0304 | 39.78 |
| DepthSame/100 | 0.3798 *** | 1.4620 | 10.71 |
| DepthOpp/100 | -0.9478 *** | 0.3876 | -9.27 |
| Volat*10000 | 0.0134 *** | 1.0135 | 22.57 |
| Intercept | -4.0663 *** | | -183.04 |



Electronic copy available at: https://ssrn.com/abstract=3074049

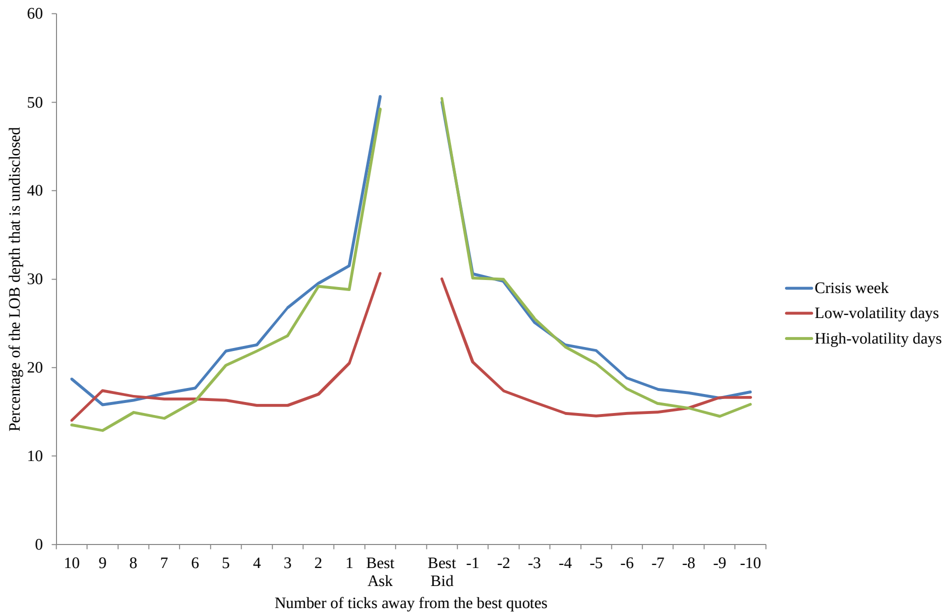
## Table XI
## Volatility anticipation strategies using HLOs

We present regressions of quote midpoint realized volatility (post order submission) on order attributes and their interactions. We provide results on two types of volatilities – volatility 30 second post order and volatility 60 second post order. The models are estimated on pooled regression basis, controlling for day, stock and time of the day fixed effects. The estimation sample for this table consists of the 30 largest stocks (in which HFTs are reasonably active) from our main sample of 100 stocks listed on the NSE of India and the sample period is December 2013. ***, **, * indicate significance at the 1%, 5%, and 10% level respectively.

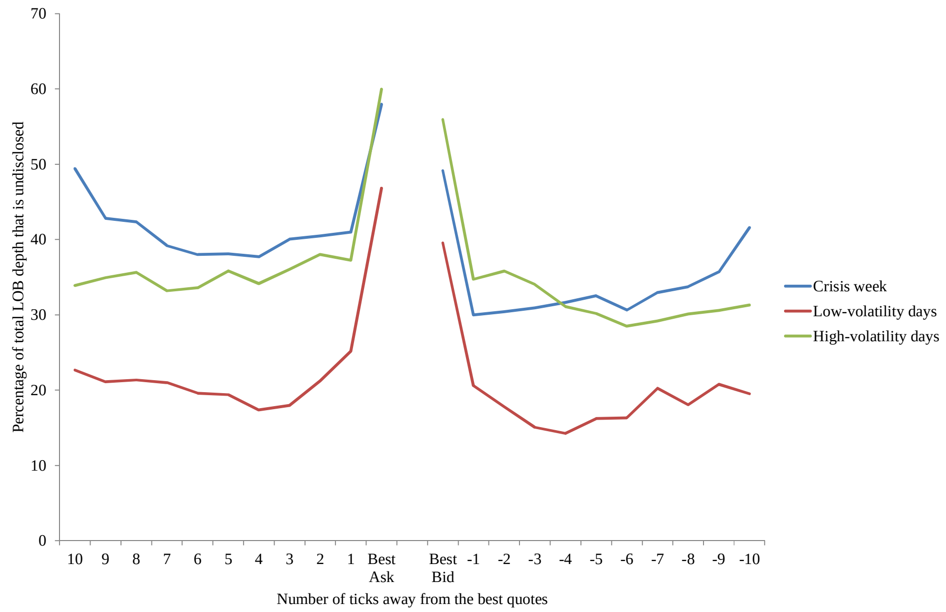
| | RVolat | |
|---|---|---|
| Coeficient | 30 sec. | 60 sec. |
| Intercept | 0.7609 *** | 1.3794 *** |
| LagRVolat | 0.3160 *** | 0.3712 *** |
| HLO | -0.3735 *** | -0.5784 *** |
| NAGR | 0.0268 *** | 0.0048 |
| HLONAGR | 0.1841 *** | 0.3136 *** |
| HFT | 0.4242 *** | 0.5355 *** |
| HLOHFT | 0.7382 *** | 0.9701 *** |
| NAGRHFT | -0.3944 *** | -0.4989 *** |
| HLONAGRHFT | 1.9011 *** | 2.5756 *** |
| AAT | 0.0451 *** | 0.0198 *** |
| HLOAAT | 0.2661 *** | 0.4428 *** |
| NAGRAAT | -0.0325 *** | -0.0200 ** |
| HLONAGRAAT | -0.1523 *** | -0.2741 *** |
| Stock fixed effects | Yes | Yes |
| Daily fixed effects | Yes | Yes |
| Intraday fixed effects | Yes | Yes |
| Obs. | 16459038 | 16406676 |
| Adj. $R^2$ | 0.2255 | 0.3128 |



Electronic copy available at: https://ssrn.com/abstract=3074049




(a) Large Cap stocks




(b) Small Cap stocks

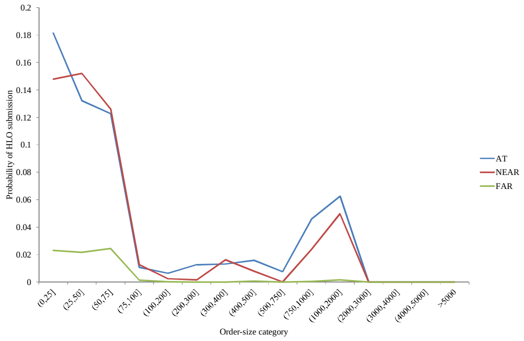
**Figure 1**

**Hidden depth in the Nasdaq Limit order book**

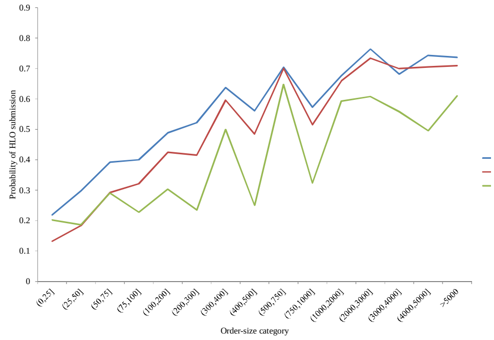
We plot the cross-sectional average percentage of the hidden depth in the Nasdaq limit order book for the "crisis week" of September 2008, the five days (per stock) with lowest and highest volatility. While the study sample consists of 120 firms split into large, mid-sized and small caps, we show results here for the large and small cap samples. We use order by order data collected from one-minute snapshots of the ten best ask and bid order book levels.



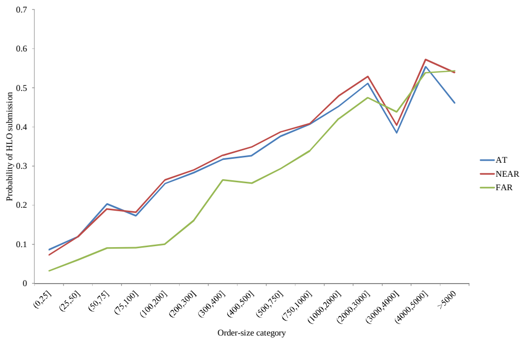
Electronic copy available at: https://ssrn.com/abstract=3074049




2.a. HFTs' probability of HLO submission




2.b. AATs' probability of HLO submission




2.c. NATs' probability of HLO submission

**Figure 2**

**Probability of submitting an HLO conditional on order size and aggressiveness**

We plot estimated cross-sectional daily average probabilities of hidden limit order (HLO) submission in the NSE conditional on order size and order aggressiveness. We distinguish between high frequency traders (HFTs), agency algorithmic traders (AATs) and non-algorithmic traders (NATs). The sample consists of the 30 largest stocks from our size-stratified sample of 100 stocks listed on the NSE between October and December 2013. We combine the limit order book levels into three groups: at the best quotes ("At"); from the best quotes up to 5 ticks away ("Near"), and the rest ("Far"). For each order size, level of aggressiveness, and type of trader, we provide the percentage of HLOs out of all the non-marketable limit orders submitted. Figures 1.a, 1.b, and 1.c provide the findings for HFTs, AATs, and NATs respectively.



Electronic copy available at: https://ssrn.com/abstract=3074049

# Order Exposure in High Frequency Markets

## Internet Appendices



Electronic copy available at: https://ssrn.com/abstract=3074049

## Appendix A
## Building the Limit Order Book of from NSE order data

The National Stock Exchange of India (NSE) provides two types of files: order files and trade files. Order files contain all the message traffic. Each order has an identification code that allows us to follow the history of the order from submission to execution/cancellation/expiry. Messages are time-stamped to the nearest microsecond. For each order, we know the type of message (new submission, revision, or cancellation); the type of order (limit order or market orders); the type of trader submitting the order (HFT, AAT or NAT, based on the identification of trader accounts described in the accompanying paper); the order direction (buy or sell), and the total size of the order. For limit orders we also know the limit price, and for iceberg orders we know both the total size and the displayed size (hidden volume is the difference between the total size and the displayed size). The file also identifies orders with special conditions: immediate-or-cancel, and on stop. The trade files provide, for each trade, the buy and sell orders matched; the type of trader submitting each order; the trade size, and the trade price.

We start each day assuming the limit order book (LOB) is empty. We use the registers of the opening auction from the order file to build the LOB pre-allocation. Orders in the NSE are sorted by price-time priority, with market orders having priority over limit orders, no matter the time of submission. The trade files provide the information about orders matched at the allocation price of the opening call auction. Non-allocated market orders at the end of the auction time are transformed into limit orders at the allocation price. If there are no trade registers associated with the opening auction it indicates that there was no allocation price. In such cases, market orders are stored at the closing price of the previous session. The result is the initial snapshot of the LOB for the corresponding day.

Then, we update the state of the LOB conditioned on each and every posterior message (new submission, revision, or cancellation) during the continuous session. We match the order and the trade files, checking that every market order and every marketable limit order submitted have their corresponding trade registers. By doing so, we can also discern the actual direction of each trade, i.e., whether the trade is buyer- or seller-initiated.



Electronic copy available at: https://ssrn.com/abstract=3074049

During the continuous session HLOs orders are allowed. As in other markets around the world, the hidden part of the iceberg orders loses time priority against displayed limit orders. Accordingly, every time the displayed volume unit of an iceberg order is exhausted, the emerging new displayed volume unit moves to the end of the queue of all standing limit orders at the same price. Our program allows us to obtain snapshots of both the displayed and the hidden components of the LOB at every instant during the continuous session.

Order revisions are the most common type of message in the NSE. These revisions can change the order size, the limit price, or both. Some of these updates can change the priority of the execution of the order. In particular, increases in volume will cause losing time priority. Decreases in volume, however, will not change priority. Obviously, increases (decreases) in the limit price of a standing limit order to buy (sell) will increase price priority. Changes in hidden volume with no change in displayed volume are possible. In that case, the displayed part of the iceberg order does not lose time priority. We update the state of the LOB after each revision to reflect these changes in price-time priority.

Changes in the type of order, from “on stop” to ordinary or the other way around are possible, but not very frequent. When an “on stop” order changes to ordinary order, it is treated as a new submission. When an ordinary order changes to “on stop”, it is removed from the LOB. Orders on stop can be revised while not activated. Once activated, a new register indicates the change in status and the final conditions under which the order reaches the LOB. At that point, the order is treated as an ordinary new submission. Immediate-or-cancel orders only change the book if executed and, therefore, generate a trade.

The best proof that our program works is that the resulting LOB file and the trade file perfectly match. When a marketable limit or a market order is submitted, the associated sequence of trade registers is consistent with what can be inferred by matching the incoming aggressive order with the price-time priority sorted orders standing in the LOB, and controlling for hidden volume. Additionally, there are no inconsistencies between the timing of order flow events and the timing of the associated trades.



Electronic copy available at: https://ssrn.com/abstract=3074049

Electronic copy available at: https://ssrn.com/abstract=3074049

## Appendix B

### Table IB I
### Nasdaq descriptive statistics

We provide evidence on the use of hidden limit orders (HLOs) by HFTs in the US using the Nasdaq's HFT database. We provide results for the whole sample except the crisis week of September 2008. The analysis is based on order by order data collected from one-minute snapshots of the 10 best ask and bid LOB levels. We distinguish between two types of traders: high-frequency traders (HFTs) and non-HFTs ("Others"). The database includes a HFT "flag" that identifies the HFT orders in the LOB. It also includes a "HLO" flag that allows us to distinguish HLOs from DLOs. The sample consists of 120 Nasdaq-listed firms that we split into 3 subsamples of 40 stocks based on market capitalization. Tests are based on the non-parametric rank sum test of Wilcoxon (1963).

| Cross-sectional daily average statistic | All | Large | Mid | Small |
|---|---|---|---|---|
| Displayed relative spread (bsp) | 31.99 | 7.40 | 22.57 *** | 65.99 *** |
| Posted relative spread (bsp) | 26.27 | 6.31 | 19.07 *** | 53.41 *** |
| Displayed depth at the best quotes ($US) | 62621.35 | 153819.71 | 28198.41 *** | 5845.91 *** |
| Total depth at the best quotes ($US) | 96453.74 | 224053.26 | 49520.39 *** | 15787.58 *** |
| Displayed depth 5 best quotes ($US) | 445666.30 | 1136466.80 | 157180.80 *** | 43351.34 *** |
| Total depth 5 best quotes ($US) | 555882.43 | 1362283.70 | 224485.33 *** | 80878.30 *** |
| Hidden volume at the best quotes (% time) | 71.72 | 78.95 | 69.30 *** | 66.91 *** |
| Hidden volume within the best displayed quotes (% time) | 72.08 | 48.99 | 77.75 *** | 89.50 *** |
| HFTs at the best quotes (% time) | 67.41 | 93.23 | 62.89 *** | 46.11 *** |
| HFTs contribution to the best quotes depth (%) | 30.90 | 50.41 | 25.36 *** | 16.92 *** |
| HFTs contribution to the 5 best quotes depth (%) | 26.80 | 36.84 | 23.58 *** | 19.97 *** |

***, **, * means statistically different than large caps



Electronic copy available at: https://ssrn.com/abstract=3074049

**Table IB II**
**Use of HLOs in the Nasdaq market**

We provide daily cross-sectional average statistics on the use of undisclosed limit orders (HLOs) and disclosed limit orders (DLOs) in the Nasdaq. The analysis is based on order by order data collected from one-minute snapshots of the 10 best ask and bid LOB levels. We distinguish between two types of traders: high-frequency traders (HFTs) and non-HFTs. The sample consists of 120 NASDAQ-listed firms that we split into 3 subsamples of 40 stocks based on market capitalization. Tests are based on the non-parametric rank sum test of Wilcoxon (1963). ***, **, * indicate differences between HFTs and non-HFTs are significant at the 1%, 5%, and 10% level respectively.

| | | Trader types | | | | | |
|---|---|---|---|---|---|---|---|
| | | HFTs | | non-HFTs | | Dif. | |
| Variable | | Orders | Vol./1000 | Orders | Vol./1000 | Orders | Vol./1000 |
| HLOs | Large | 1223.60 | 624.24 | 1238.25 | 5252.19 | -14.65 *** | -4627.95 *** |
| | Mid | 624.07 | 422.34 | 1089.13 | 1592.48 | -465.06 *** | -1170.14 *** |
| | Small | 634.34 | 700.77 | 1348.82 | 1654.52 | -714.48 *** | -953.75 *** |
| | | Orders (%) | Vol. (%) | Orders (%) | Vol. (%) | Orders (%) | Vol. (%) |
| % HLOs/LOs | Large | 21.80 | 15.25 | 15.39 | 21.74 | 6.40 *** | -6.49 ** |
| | Mid | 23.17 | 34.71 | 15.43 | 27.13 | 7.74 *** | 7.58 *** |
| | Small | 31.65 | 47.84 | 19.74 | 35.31 | 11.90 *** | 12.54 *** |
| Share of HLOs | Large | 44.00 | 16.83 | 56.00 | 83.17 | -12.01 *** | -66.34 *** |
| | Mid | 33.33 | 28.04 | 66.67 | 71.96 | -33.34 *** | -43.92 *** |
| | Small | 30.97 | 34.46 | 69.03 | 65.54 | -38.07 *** | -31.08 *** |



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table IB III
## Hidden order and volume placement in the order book in the Nasdaq market

We examine the placement of hidden limit orders and displayed limit orders across 3 levels of the limit order book (LOB) - (a) at the best quotes ("At"), (b) from the best quotes up to 5 ticks away ("Near"), and (c) the rest ("Far"). Panel A includes the results on share volume of orders and Panel B presents the results on number of orders. We provide results for the whole sample period except the crisis week of September 2008. The analysis is based on order by order data collected from one-minute snapshots of the 10 best ask and bid LOB levels. We distinguish between two types of traders: high-frequency traders (HFTs) and non-HFTs ("Others"). The database includes a HFT "flag" that identifies the HFT orders in the LOB. It also includes a "HLO" flag that allows us to distinguish HLOs from DLOs. The sample consists of 120 Nasdaq-listed firms that we split into 3 subsamples of 40 stocks based on market capitalization. We only report our findings for the 40 largest. Tests are based on the non-parametric rank sum test of Wilcoxon (1963). ***, **, * indicate significance at the 1%, 5%, and 10% level respectively.

Panel A: Volume

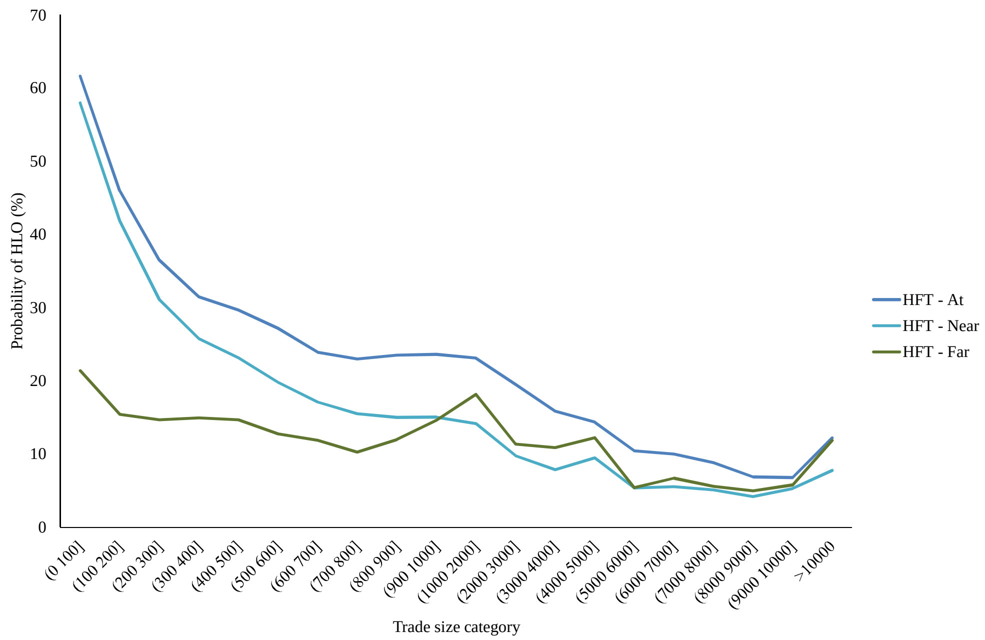
| | Trader types | | | |
|---|---|---|---|---|
| | HFTs | | Others | |
| Position | Displayed | Hidden | Displayed | Hidden |
| At | 11.48 | 32.92 | 4.96 *** | 23.70 *** |
| Near | 49.49 | 52.44 | 39.35 *** | 36.17 *** |
| Far | 39.03 | 14.64 | 55.69 *** | 40.13 *** |

Panel B: Orders

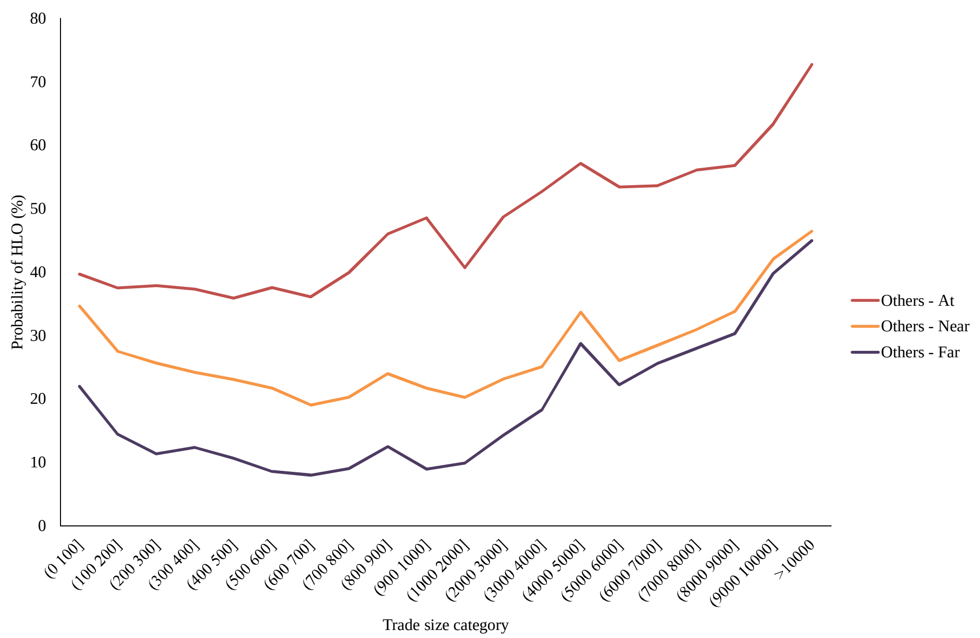
| | DLOs | HLOs | DLOs | HLOs |
|---|---|---|---|---|
| At | 9.63 | 30.26 | 8.13 *** | 22.45 *** |
| Near | 44.74 | 59.13 | 37.66 *** | 47.21 *** |
| Far | 45.63 | 10.61 | 54.21 *** | 30.34 *** |



Electronic copy available at: https://ssrn.com/abstract=3074049




IB 1a. HFTs – Large Caps




IB 1b. non-HFTs – Large Caps

**Figure IB 1. Probability of submitting an HLO conditional on order size and aggressiveness**

We plot cross-sectional daily average probability of HLO submission conditional on the order size and aggressiveness. We distinguish between high-frequency traders (HFTs) and non-HFTs (“Others”). The sample period consists of 50 non-consecutive days from 2008 to 2010. We exclude the “crisis week” of September 2008. The sample consists of 120 Nasdaq-listed firms split into three equally-sized subsamples: large, mid-sized (not reported), and small caps. We use order by order data collected from one-minute snapshots of the ten best ask and bid LOB levels. We define three levels of aggressiveness: (a) at the best quotes (“At”); (b) up to five ticks away from the best quotes (“Near”), 6 ticks away or more from the best quote (“Far”). We consider order sizes from 1 to 1,000 shares in increments of one lot (i.e., (0, 100], (100, 200] … (900 1,000]), from 1,000 to 10,000 shares in increments of 10 lots (i.e., (1,000 2,000], (2,000 3,000] … (9,000 10,000]), and orders of size greater than 10,000 shares



Electronic copy available at: https://ssrn.com/abstract=3074049

# Appendix C
# List of variables and scaling factors

| Variable | Definition | Multiplier |
|---|---|---|
| AAT | Indicator variable that equals 1 for orders submitted by AATs and 0 otherwise | |
| Aggr | Distance of the order's limit price from the opposite quote price, suitably signed (a higher value indicates a more aggressively priced order) divided by the quote midpoint | 100 |
| AggrUnd | Distance of the undercutted order's limit price from the opposite quote price, suitably signed (a higher value indicates a more aggressively priced order) divided by the quote midpoint | |
| Buy | Indicator variable that equals 1 for buy orders and 0 otherwise | |
| DispSizeUnd | Displayed size of the undercutted order | |
| DepthOpp | Displayed depth at the best ask (bid) for a buy (sell) order divided by the average daily trading volume | 1/10000000 |
| DepthSame | Displayed depth at the best bid (ask) for a buy (sell) order divided by the average daily trading volume | 1000 |
| HFT | Indicator variable that equals 1 for orders submitted by HFT and 0 otherwise | |
| HidVol | Indicator equals to 1 if presence of hidden volume on the same side is deteced, 0 otherwise | |
| HLO | Indicator variable that equals 1 for hidden orders and 0 otherwise | |
| LagRVolat | First lag of RVolat | |
| LastBuy | Indicator variable that equals 1 If the last trade is buyer initiated and 0 otherwise | |
| LastHalfHour | Indicator variable that equals 1 for orders submitted in the last hour of the trading day and 0 otherwise | |
| LOBImb | Percentual difference between the accumulated displayed depth in the best five bid and ask quotes of the book, suitably signed (i.e., positive when same side depth exceeds opposite side depth) | 100 |
| DistMidQ | Difference between the quote midpoint and the limit price of the order | 100 |
| Mom | Continuously compound quote midpoint return in last 5 minutes | |
| NAT | Indicator variable that equals 1 for orders submitted by NATs and 0 otherwise | |
| OI | Buyer-initiated volume minus seller-initiated volume divided by total volume in last 5 minutes | |
| OrdSize | Total (displayed plus hidden) size of the order divided by average daily trading volume | |
| NonAggr | Indicator variable that equals 1 for orders submitted beyond the prevailing best quotes and 0 otherwise | |
| RVolat | Realized volatility computed from quote midpoints collected at regular 1-second intervals | |
| RelTrdFreq | Number of shares traded in last 30 minutes divided by number of shares traded in last 60 minutes | |
| RSpr | Bid-ask spread divided by the quote midpoint | |
| TrdFreq | Number of shares traded per second within the last $k$ minutes | |
| Volat | Sum of the squared continuously compound quote-midpoint return over the last $k$ minutes | 1/1000000 |

$k$ = 60 minutes in Table IV and VI, 5 minutes in Tables V, VII, and XI



Electronic copy available at: https://ssrn.com/abstract=3074049

## Appendix D

## Additional result and Methodological details

### Table IA IV
### The order exposure decision

We study the determinants of the order (non-) exposure decision of high frequency traders (HFTs), agency algorithmic traders (AATs) and non-algorithmic traders (NATs). We use logistic models (Panel A) of order characteristics and market conditions to study the choice between submitting a hidden limit order (HLO) and a fully displayed limit order (DLO). We exclude all market and marketable limit orders. The dependent variable equals one (zero) if the NAT submits a HLO (DLO). We use Tobit models (Panel B) of order characteristics and market conditions to study the decision of how much volume of a limit order is hidden. The dependent variable here is the amount of shares hidden, normalized by the stock's average daily trading volume. The models are estimated on a stock-by-stock basis, and we report aggregated coefficients and t-statistics using the approach in Chordia, Roll, and Subrahmanyam (2005). The estimation sample for this table consists of the 30 largest stocks (in which HFTs are reasonably active) from our main sample of 100 stocks listed on the NSE. The sample period is December 2013. ***, **, * indicate significance at the 1%, 5%, and 10% level respectively.

| Panel A: Decision to hide - logistic model | | | |
|---|---|---|---|
| | ATs | | |
| Variable | HFTs | AATs | NATs |
| Intercept | -3.9108 *** | -0.8195 ** | -1.8061 *** |
| Price aggressiveness | 2865.7587 *** | 511.3416 *** | 65.7729 *** |
| Total order size | 31.7138 ** | 19.9858 *** | 18.3290 *** |
| Relative spread | 1558.2250 *** | -69.7108 | -4.0103 |
| Depth same side | -586.9779 *** | -216.5916 *** | -88.1710 *** |
| Depth opposite side | 39.8854 | 50.2558 *** | -30.9239 ** |
| Stock volatility | -0.0141 | -0.0031 | -0.0062 *** |
| Waiting time | -50.3939 * | 24.9165 | 15.5722 ** |
| Trade frequency | -1.5337 | -0.4582 | -0.7669 ** |
| Hidden same side | -3.0559 | 0.0679 | -0.2246 |
| LOB order imbalance | 15.7592 | 0.4677 | -0.2394 |
| Last trade size | -3.4383 *** | -2.0167 ** | -0.4277 * |
| Market volatility | -0.0017 * | -0.0014 | -0.0001 |
| Last half hour indicator | 572.6601 *** | 72.4503 | -169.1852 *** |
| **Panel B: Magnitude of hidden volume - Tobit model** | | | |
| Intercept | -0.0041 | -0.0007 ** | -0.0031 *** |
| Price aggressiveness | 0.2880 *** | 0.0726 *** | 0.0607 ** |
| Total order size | 0.0043 | 0.0067 *** | 0.0055 *** |
| Relative spread | 0.1933 *** | -0.0168 | 0.0709 |
| Depth same side | -0.0479 ** | -0.0461 | -0.0332 * |
| Depth opposite side | 0.0051 | 0.0035 | -0.0278 ** |
| Stock volatility | 0.5508 | 0.0208 | -0.0501 *** |
| Waiting time | -0.0060 | 0.0014 | 0.0049 *** |
| Trade frequency | -0.0075 | -0.0002 | -0.0005 |
| Hidden same side | -0.0874 | -0.0033 | 0.0004 |
| LOB order imbalance | 0.0007 | 0.0010 | 0.0016 |
| Last trade size | -0.0003 | -0.0003 | -0.0001 |
| Market volatility | 0.0000 | 0.0000 | 0.0000 |
| Last half hour indicator | 0.0544 | 0.0843 | -0.0466 |

***,**,* means statistically significant at the 1%, 5% and 10% level, respectively



Electronic copy available at: https://ssrn.com/abstract=3074049

**Table VIII (Panels A and B)**
**Impulse-response function (VAR model)**

We investigate the permanent price impact (informational content) of different types of orders by different trader types. As order types, we consider market and marketable limit orders (Trades), displayed non-marketable limit orders (DLO), non-marketable hidden orders (HLO), and cancellations of standing limit orders. We consider three types of traders: HFTs, AATs, and NATs. To estimate the permanent price impact of each type of order, we use the VAR approach of Hasbrouck (1991a), as extended by Fleming, Mizrach, and Nguyen (2015) and Brogaard, Hendershott, and Riordan (2019). The model is defined in event time, where each order is an observation ($t$), and estimated per stock-day. We assume the trading process restarts each day, resetting all lagged values to zero. The model is,

$$\begin{aligned}
r_t &= \sum_{j=1}^{n}\alpha_j^0 r_{t-j} + \sum_{j=0}^{n}\beta_j^{0,1}X_{t-j}^1 + \sum_{j=0}^{n}\beta_j^{0,2}X_{t-j}^2 + \cdots + \sum_{j=0}^{n}\beta_j^{0,12}X_{t-j}^{12} + \varepsilon_t \\
X_t^1 &= \sum_{j=1}^{n}\alpha_j^1 r_{t-j} + \sum_{j=1}^{n}\beta_j^{1,1}X_{t-j}^1 + \sum_{j=1}^{n}\beta_j^{1,2}X_{t-j}^2 + \cdots + \sum_{j=1}^{n}\beta_j^{1,12}X_{t-j}^{12} + \mu_t^1 \\
X_t^2 &= \sum_{j=1}^{n}\alpha_j^2 r_{t-j} + \sum_{j=1}^{n}\beta_j^{2,1}X_{t-j}^1 + \sum_{j=1}^{n}\beta_j^{2,2}X_{t-j}^2 + \cdots + \sum_{j=1}^{n}\beta_j^{2,12}X_{t-j}^{12} + \mu_t^2 \qquad [1] \\
\vdots \;&= \qquad\qquad\qquad \vdots \\
X_t^{12} &= \sum_{j=1}^{n}\alpha_j^{12} r_{t-j} + \sum_{j=1}^{n}\beta_j^{12,1}X_{t-j}^1 + \sum_{j=1}^{n}\beta_j^{12,2}X_{t-j}^2 + \cdots + \sum_{j=1}^{n}\beta_j^{12,12}X_{t-j}^{12} + \mu_t^{12}
\end{aligned}$$

or in compact form

$$A_0 y_t = \sum_{j=1}^{n} A_j y_{t-j} + \xi_t$$

where $y_t' = (r_t, X_t)$ is the 1x13 vector of contemporaneous dependent variables and $\xi_t' = (\varepsilon_t, \mu_t)$ is the 1x13 vector of innovations to the dependent variables; $r_t$ is the continuously compound quote midpoint return expressed in basis points; $X_t$ is a vector of order-flow related variables. By combining the 3 types of traders and the 4 types of events/orders, we have 12 possible order-flow categories ($X^1$ *to* $X^{12}$): NAT/AT/HFT – Trade, NAT/AT/HFT – DLO, NAT/AT/HFT – HLO, NAT/AT/HFT – Cancellation.



Electronic copy available at: https://ssrn.com/abstract=3074049

Each $X^k$ can take one of three possible values: 0, 1 or -1. $X^k$ = 1(-1) if order $t$ is a buy (sell) order of type $k$ and zero otherwise. Because the model is defined in event time, whenever $X^k = 1$, $X^z = 0\ \forall z \neq k$. The number of lags ($n$) is stock-day specific and determined using the Schwarz' Bayesian Information Criterion (SBIC), which Lütkepohl (2005, p. 148-152) shows provides consistent estimates of the true lag order.

As in Hasbrouck (1991a) original VAR model, we assume contemporaneous causality running from the order flow to the changes in prices. Accordingly, the 13x13 matrix $A_0$ equals

$$A_0 = \begin{pmatrix} 1 & -\beta_0^{0,1} & -\beta_0^{0,2} & -\beta_0^{0,3} & \cdots & -\beta_0^{0,12} \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

and the variance-covariance matrix of the residuals becomes,

$$Var(\xi_t) = \begin{pmatrix} \sigma_\varepsilon^2 & \overline{0} \\ \overline{0}' & \Omega \end{pmatrix}$$

where $\sigma_\varepsilon^2$ is the variance of the innovation to $r$; $\overline{0}$ is a 1x12 vector of zeros, and $\Omega$ is the variance-covariance matrix of $X$.

Because the model is defined in event time, $\Omega$ is near-diagonal. For a representative stock-day, the average contemporaneous correlation across $\mu^k$ innovations is about 0.1%. Nonetheless, we follow Brogaard et al. (2019) in computing orthogonalized and order independent IRFs. The IRF for trades and orders (cancellations) is computed for a unitary positive (negative) shock at period $t$ = 0, assuming the model is in a steady state (i.e., all lagged variables equal to zero), and the subsequent price impact (in basis points) is accumulated over the next 20 periods. The model is estimated for each stock-day, and we report the average IRF across stock-days. Statistical significance is clustered by stock and day (e.g., Thompson, 2011).



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table VIII (Panel C)
## Order Flow-related efficient variance decomposition

The analysis summarized in Table VIII (Panels A and B) provides estimates of the average information content of particular types of orders submitted by particular types of traders. That analysis, however, should not be interpreted in terms of overall contributions to price discovery, since we ignore the frequency with which each event takes place. In Table VIII Panel C, we follow Hasbrouck (1991b) to estimate the relative contribution of the different (trader type, order type) binomials to the component of the long-run variance of the stock attributable to the order flow.

From the VMA representation of a VAR model similar to [1] but with only two variables (quote midpoint changes and trades), Hasbrouck (1991b) obtains an estimate of the long-run variance of the corresponding asset, say Σ, which he further decomposes into a trade-related component, due to the innovations to the trading process (i.e., $\mu$ in [2]) and a trade-unrelated component, due to the innovations to the quote midpoint changes (i.e., $\varepsilon$ in [1]). Because of the one-directional causality assumption (from trades to quotes), ε are contemporaneously uncorrelated with $\mu$.

Specifically, the VAR model in [1] can be re-written as

$$A_0 y_t = \sum_{j=1}^{n} A_j y_{t-j} + \varepsilon_t \rightarrow A(L) y_t = \varepsilon_t \, .$$

where $L$ is the lag operator, that is, $L^j y_t = y_{t-j}$, and *A(L)* is a lag polynomial, that is, $A(L) = A_0 - \sum_{j=1}^{n} A_j L^j$. Its VMA representation would be

$$y_t = \Psi(L)\xi_t \qquad [2]$$

In expanded form, the VMA in [2] is as follows,



Electronic copy available at: https://ssrn.com/abstract=3074049

$$
\begin{aligned}
r_t &= \sum_{j=1}^{\infty}\theta_j^0\varepsilon_{t-j} + \sum_{j=0}^{\infty}\phi_j^{0,1}\mu_{t-j}^1 + \sum_{j=0}^{n}\phi_j^{0,2}\mu_{t-j}^2 + \cdots + \sum_{j=0}^{\infty}\phi_j^{0,12}\mu_{t-j}^{12} + \varepsilon_t \\
X_t^1 &= \sum_{j=1}^{\infty}\theta_j^1\varepsilon_{t-j} + \sum_{j=1}^{\infty}\phi_j^{1,1}\mu_{t-j}^1 + \sum_{j=1}^{\infty}\phi_j^{1,2}\mu_{t-j}^2 + \cdots + \sum_{j=1}^{\infty}\phi_j^{1,12}\mu_{t-j}^{12} + \mu_t^1 \\
X_t^2 &= \sum_{j=1}^{\infty}\theta_j^2\varepsilon_{t-j} + \sum_{j=1}^{\infty}\phi_j^{2,1}\mu_{t-j}^1 + \sum_{j=1}^{\infty}\phi_j^{2,2}\mu_{t-j}^2 + \cdots + \sum_{j=1}^{\infty}\phi_j^{2,12}\mu_{t-j}^{12} + \mu_t^2 \\
\vdots &= \qquad\qquad\qquad \vdots \\
X_t^{12} &= \sum_{j=1}^{\infty}\theta_j^{12}\varepsilon_{t-j} + \sum_{j=1}^{\infty}\phi_j^{12,1}\mu_{t-j}^1 + \sum_{j=1}^{\infty}\phi_j^{12,2}\mu_{t-j}^2 + \cdots + \sum_{j=1}^{\infty}\phi_j^{12,12}\mu_{t-j}^{12} + \mu_t^{12}
\end{aligned} \qquad [3]
$$

Notice that [3] keeps the contemporaneous causality flow from orders to trades.

Hasbrouck (1991b) defines the order-flow related component of the long-run variance of the stock ($I_x$) as

$$I_x = \frac{Var\left(E\left[\Delta m_t \mid \mu_t\right]\right)}{Var(\Delta m_t)} = \frac{\sigma^2_{\Delta m,\mu}}{\sigma^2_{\Delta m}} \qquad [4]$$

Now assume that any non-diagonal element in $Var(\mu_t) = \Omega$ is negligible; they actually are, as we have explained before. Let $\phi_j^0 = \left(\phi_j^{0,1}, \phi_j^{0,2}, \ldots, \phi_j^{0,12}\right)$ be the row vector of order flow related coefficients at lag $j$ in the $r_t$ equation of the VMA model [3]. Therefore, the row vector of cumulated impacts of order-flow-related unitary shocks is $\phi = \sum_{j=0}^{\infty}\phi_j^0 = \left(\sum_{j=0}^{\infty}\phi_j^{0,1}, \sum_{j=0}^{\infty}\phi_j^{0,2}, \ldots, \sum_{j=0}^{\infty}\phi_j^{0,12}\right)$. Similarly, the cumulated impact of a unitary order flow unrelated shock is $\theta = 1 + \sum_{j=1}^{\infty}\theta_j^0$. Hasbrouck (1991b) shows that the long run variance (the variance of the efficient price) can be computed from the VMA coefficients as $\sigma^2_{\Delta m} = \phi\Omega\phi' + \theta^2\sigma^2_{\varepsilon}$, and the order flow related efficient variance can be decomposed as

$$\sigma^2_{\Delta m,\mu} = \phi\Omega\phi' = \left(\sum_{j=0}^{\infty}\phi_j^{0,1}\right)^2\sigma^2_{\mu_1} + \left(\sum_{j=0}^{\infty}\phi_j^{0,2}\right)^2\sigma^2_{\mu_2} + \ldots + \left(\sum_{j=0}^{\infty}\phi_j^{0,12}\right)^2\sigma^2_{\mu_{12}} \quad , [5]$$

that is, the sum of the variance of the IRFs of order flow related shocks.

In Table VIII Panel C, we provide the contribution of each (trader type, order type) binomial's related shocks to the long-run variance component in eq. [5].



Electronic copy available at: https://ssrn.com/abstract=3074049

## Table IX
## Information shares

Following Hasbrouck (1995), we estimate stock-day information shares (IS) for different trader-type/order-type combinations. In contrast to analyses run in event time, Hasbrouck's IS approach evaluates price discovery using trader-type/order-type specific quotes collected at regular time intervals. As noted in Brogaard et al. (2019), event time analyses do not account for tiny differences in the response of different traders to new public information releases, which result in the subsequent price discovery being attributed to the fastest traders. Thus, the IS uses a more conservative timing approach to price discovery.

We compute trader-type/order-type specific quote midpoints prevailing at the end of each second. We consider three types of traders (HFTs, AATs, and NATs) and two types of orders (DLOs and HLOs). For each trader type, we collect the best ask and bid quotes supported by standing DLOs and compute the quote midpoint by averaging the best ask and bid quotes. In case there are no DLOs standing on the LOB for that trader type, the observation is replaced by the closest preceding non-missing observation. For HLOs, we proceed in the same way.

Hasbrouck's (1995) approach decomposes the variance of the underlying efficient price into components attributable to the different trader-type/order type pairs, the so-called "information shares". The first step of this methodology estimates a Vector Error-Correction (VEC) Model for each stock-day, under the assumption that the quote midpoints are co-integrated. The VEC model is reported in eq. [6]

$$
\begin{aligned}
\Delta q_t^{HI} &= \alpha^{HI\prime}\beta q_{t-1} + \sum_{j=1}^{n}\phi_j^{HI,HI}\Delta q_t^{HI} + \sum_{j=1}^{n}\phi_j^{HI,HD}\Delta q_t^{HD} + \sum_{j=1}^{n}\phi_j^{HI,AI}\Delta q_t^{AI} + \sum_{j=1}^{n}\phi_j^{HI,AD}\Delta q_t^{AD} + \sum_{j=1}^{n}\phi_j^{HI,NI}\Delta q_t^{NI} + \sum_{j=1}^{n}\phi_j^{HI,ND}\Delta q_t^{ND} + \varepsilon_t^{HI} \\
\Delta q_t^{HD} &= \alpha^{HD\prime}\beta q_{t-1} + \sum_{j=1}^{n}\phi_j^{HD,HI}\Delta q_t^{HI} + \sum_{j=1}^{n}\phi_j^{HD,HD}\Delta q_t^{HD} + \sum_{j=1}^{n}\phi_j^{HD,AI} r_t^{AI} + \sum_{j=1}^{n}\phi_j^{HD,AD}\Delta q_t^{AD} + \sum_{j=1}^{n}\phi_j^{HD,NI}\Delta q_t^{NI} + \sum_{j=1}^{n}\phi_j^{HD,ND}\Delta q_t^{ND} + \varepsilon_t^{HD} \\
\Delta q_t^{AI} &= \alpha^{AI\prime}\beta q_{t-1} + \sum_{j=1}^{n}\phi_j^{AI,HI}\Delta q_t^{HI} + \sum_{j=1}^{n}\phi_j^{AI,HD}\Delta q_t^{HD} + \sum_{j=1}^{n}\phi_j^{AI,AI}\Delta q_t^{AI} + \sum_{j=1}^{n}\phi_j^{AI,AD}\Delta q_t^{AD} + \sum_{j=1}^{n}\phi_j^{AI,NI}\Delta q_t^{NI} + \sum_{j=1}^{n}\phi_j^{AI,ND}\Delta q_t^{ND} + \varepsilon_t^{AI} \\
\Delta q_t^{AD} &= \alpha^{AD\prime}\beta q_{t-1} + \sum_{j=1}^{n}\phi_j^{AD,HI}\Delta q_t^{HI} + \sum_{j=1}^{n}\phi_j^{AD,HD}\Delta q_t^{HD} + \sum_{j=1}^{n}\phi_j^{AD,AI}\Delta q_t^{AI} + \sum_{j=1}^{n}\phi_j^{AD,AD}\Delta q_t^{AD} + \sum_{j=1}^{n}\phi_j^{AD,NI}\Delta q_t^{NI} + \sum_{j=1}^{n}\phi_j^{AD,ND}\Delta q_t^{ND} + \varepsilon_t^{AD} \\
\Delta q_t^{NI} &= \alpha^{NI\prime}\beta q_{t-1} + \sum_{j=1}^{n}\phi_j^{NI,HI}\Delta q_t^{HI} + \sum_{j=1}^{n}\phi_j^{NI,HD}\Delta q_t^{HD} + \sum_{j=1}^{n}\phi_j^{NI,AI}\Delta q_t^{AI} + \sum_{j=1}^{n}\phi_j^{NI,AD}\Delta q_t^{AD} + \sum_{j=1}^{n}\phi_j^{NI,NI}\Delta q_t^{NI} + \sum_{j=1}^{n}\phi_j^{NI,ND}\Delta q_t^{ND} + \varepsilon_t^{NI} \\
\Delta q_t^{ND} &= \alpha^{ND\prime}\beta q_{t-1} + \sum_{j=1}^{n}\phi_j^{ND,HI}\Delta q_t^{HI} + \sum_{j=1}^{n}\phi_j^{ND,HD}\Delta q_t^{HD} + \sum_{j=1}^{n}\phi_j^{ND,AI}\Delta q_t^{AI} + \sum_{j=1}^{n}\phi_j^{ND,AD}\Delta q_t^{AD} + \sum_{j=1}^{n}\phi_j^{ND,NI}\Delta q_t^{NI} + \sum_{j=1}^{n}\phi_j^{ND,ND}\Delta q_t^{ND} + \varepsilon_t^{ND}
\end{aligned}
\qquad [6]
$$



Electronic copy available at: https://ssrn.com/abstract=3074049

where $q_t' = (q_t^{HI}, q_t^{HD}, q_t^{AI}, q_t^{AD}, q_t^{NI}, q_t^{ND})$ is the transposed quote-midpoint vector. For each stock-day we obtain the optimal lag length (*n* in eq. [6]) using the SBIC. We determine the number of linearly independent co-integration relationships (the co-integration rank) using the trade statistic proposed by Johansen (1995). Under the assumption that the difference of any two quote midpoint series in *q* is co-integrated of order (1,1), the co-integration rank should be equal to five. This is actually the case for all stock-days except for 10 cases. We exclude those abnormal stock-day observations. For the same reason, the co-integrating matrix $\beta$ should look like

$$\beta = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

We do not restrict the $\beta$ coefficients, but our estimates corroborate the above assumption about this matrix with the co-integrating vectors being the difference between two of the quote midpoints in *q*. Finally, the error-correction vector $\alpha^{j\prime} = \left(\alpha^{j,HI}, \alpha^{j,HD}, \alpha^{j,AI}, \alpha^{j,AD}, \alpha^{j,ND}, \alpha^{j,HI}\right)$, for *j* = {HI, HD, AI, AD, ND, HI} captures the sensitivity of the *j*-th quote to deviations from other trader-type/order-type quotes.

The VEC model [6] can be written in a more compact form as

$$\Delta q_t = \alpha\beta' q_{t-1} + B(L)\Delta q_{t-1} + \varepsilon_t \qquad [6']$$

The VMA representation of [6'] is

$$\Delta q_t = \Psi(L)\varepsilon_t \qquad [7]$$

Co-integration entails $\beta'\Psi(1) = 0$, with $\Psi(1) = \sum_{j=1}^{\infty} \Psi_j$ (e.g., Engle and Granger, 1987). Under the assumption in [6], Hasbrouck (1995) shows that all the rows of the impact matrix $\Psi(1)$ are identical



Electronic copy available at: https://ssrn.com/abstract=3074049

$$\Psi(1)=\begin{pmatrix}\psi_1\\ \vdots\\ \psi_6\end{pmatrix}=\begin{pmatrix}\psi\\ \vdots\\ \psi\end{pmatrix}$$

the long-run impact becomes $\psi\varepsilon_t$, and the long-run variance is

$$Var\left(\Delta m_t\right)=Var\left(\psi\varepsilon_t\right)=\psi\Omega\psi' \qquad [8]$$

To solve the identification problems that arise when the contemporaneous correlation between innovations is non-negligible, Hasbrouck (1995) suggests using the Cholesky factorization of $\Omega=FF'$, so that the IS for a given innovation is

$$IS_j=\frac{\left(\left[\psi F\right]_j\right)^2}{\psi\Omega\psi'} \qquad [9]$$

where $\left[\psi F\right]_j$ is the $j$-th element of the row vector $\psi F$.[15] The resulting factorization, however, depends on the order of the variables in the $q_t$ vector. Equation [9] will allocate a greater IS to the first quote in vector $q_t$.

Hasbrouck (1995) proposes to obtain upper and lower bounds on the IS of each quote by rotating the ordering of the variables in the $q$ vector. Unfortunately, that implies that the IS approach can only determine the contribution of each market or quote within a range. The width of this range depends on the contemporaneous correlation across quotes (e.g., Huang, 2002).

Baillie, Booth, Tse, and Zabotina (2002) and de Jong (2002) both show that the price impact vector $\psi$ and $\alpha_\perp$, the orthogonal vector of the error-correction term $\alpha_\perp\alpha'=0$, are equal up to a scale factor $\pi$, $\psi=\pi\alpha_\perp$, that drops out in the IS measure in [9]. This result largely simplifies the computation of the IS since it is not necessary to obtain the VMA representation of the VEC model. Using this result, we compute the upper and lower bounds of the IS of each trader-type/order-type pair as

$$IS_j=\frac{\left(\left[\alpha_\perp F\right]_j\right)^2}{\alpha_\perp\Omega\alpha_\perp'} \qquad [10]$$

[15] With correlated innovations the ISs are not identified since the covariance terms could be arbitrarily allocated between quotes.



Electronic copy available at: https://ssrn.com/abstract=3074049

As in former analyses, the ISs are estimated for each stock-day, and we report the average IS across stock-days. Statistical significance is clustered by stock and day (Thompson, 2011).



Electronic copy available at: https://ssrn.com/abstract=3074049